

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
26. Juni 2003 (26.06.2003)

PCT

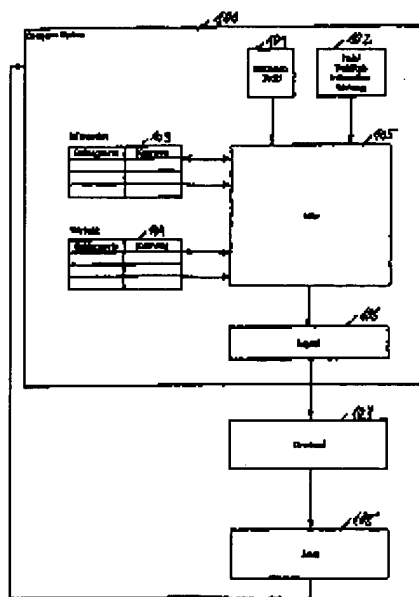
(10) Internationale Veröffentlichungsnummer  
WO 03/052648 A2

- (51) Internationale Patentklassifikation: G06F 17/60 (71) Anmelder und  
(21) Internationales Aktenzeichen: PCT/EP02/14348 (72) Erfinder: SCHÖDER, Detlef (DE/DE); Auf der  
(22) Internationales Anmeldedatum: 16. Dezember 2002 (16.12.2002) (74) Anwälte: QUERMANN, Helmut usw.; Unter den Eichen  
(25) Einreichungssprache: Deutsch 7, 65195 Wiesbaden (DE).  
(26) Veröffentlichungssprache: Deutsch (81) Bestimmungsgestaaten (national): AB, AG, AL, AM, AT,  
(30) Angaben zur Priorität: 101 61 684.8 15. Dezember 2001 (15.12.2001) DB AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,  
CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GG,  
GH, GM, HR, HU, ID, IL, IN, IS, JP, KR, KG, KP, KR,  
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,

[Fortsetzung auf der nächsten Seite]

(54) Title: SYSTEM AND METHOD FOR PRODUCTION OF A PRINTED PRODUCT INDIVIDUALISED FOR THE CUSTOMER

(54) Bezeichnung: SYSTEM UND VERFAHREN ZUR HERSTELLUNG EINES KUNDENINDIVIDUELLEN DRUCKERZEUGNISSES



101- PROFILE OF INTERESTS  
102- INFORMATION CONTENT TO ADVERTISEMENTS  
103- INFORMATION HEADLINE CODE  
104- ADVERTISEMENT HEADLINE CODE  
105- PRINT WORKS  
106- READER  
107- LAYOUT

(57) Abstract: The invention relates to a system and method for production of a printed product individualised for the customer, which offers the customer a selection of information content, depending on the variable payment (price) and the profile of interests thereof. Based on the price decision of the customer, a customer-specific relationship of optionally edited information content to customer-specific advertisement is generated. The inventive learning system carries out a dynamic matching of the profile of interests by means of a feedback mechanism. The form of the printed product is determined by the selection of the information and the advertisements and the combination thereof according to the customer-specific relationship and profile of interests by means of an inventive layout. The printed product produced as above is suitable for traditional processing by print media, customer-specific in price, quality and quantity, achieves the interests of the advertisers with targeted advertising and the interests of the producers of the printed product with regard to achievable margins by providing a highly-attractive customer-specific product. Fields of application include for example, mass-produced, individualised daily and weekly newspapers, special interest publications and general paper-based information services.

(57) Zusammenfassung: Die Erfindung betrifft ein System und Verfahren zur Herstellung eines kundenindividuellen Druckerzeugnisses, welches dem Kunden Wahlmöglichkeiten des Informationsangebotes in Abhängigkeit seines variablen Zahlungsbetrages (Preises) sowie seines Interessenprofils gibt. Dabei wird basierend auf der Preisentscheidung des Kunden ein kundenindividuelles Verhältnis von ggf. redaktionell aufbereiteter Information zu kundenspezifischer Werbung festgelegt. Das erfindungsgemäße lernende System nimmt über einen Rückkopplungsmechanismus eine dynamische Anpassung des Interessensprofils vor. Die Gestalt des Druckerzeugnisses bestimmt sich nach der Auswahl der Information und Werbung und ihrer Zusammenstellung gemäß dem kundenindividuellen Verhältnis und Interessenprofil mithilfe eines erfindungsgemässen Layout. Das erfindungsge-

WO 03/052648 A2

mäss hergestellte

[Fortsetzung auf der nächsten Seite]

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
7 September 2001 (07.09.2001)

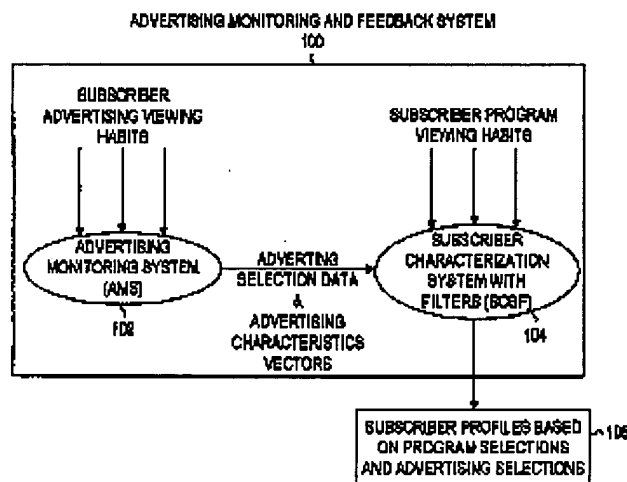
PCT

(10) International Publication Number  
**WO 01/65747 A1**

- (51) International Patent Classification<sup>7</sup>: H04H 9/00, H04N 7/025, 7/10, 7/16
- (74) Agents: RYDER, Douglas, J. et al.; 300 North Broad Street, Doylestown, PA 18901 (US).
- (21) International Application Number: PCT/US01/06459
- (22) International Filing Date: 1 March 2001 (01.03.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
09/516,983 1 March 2000 (01.03.2000) US  
09/516,314 1 March 2000 (01.03.2000) US
- (71) Applicant (for all designated States except US): EXPANSE NETWORKS, INC. [US/US]; 300 North Broad Street, Doylestown, PA 18901 (US).
- (72) Inventor; and
- (75) Inventor/Applicants (for US only): ELDERING, Charles, A. [US/US]; 214 Commons Way, Doylestown, PA 18901 (US). SYLLA, M., Lamine [SN/US]; 347 East Butler Avenue, Apt. C, New Britain, PA 18901 (US). BLASKO, John, P. [US/US]; 4 Old Mill Lane, New Hope, PA 18938 (US).
- (81) Designated States (national): AB, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GG, GH, GM, HR, HU, ID, IL, IN, IS, JP, KB, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TO).
- Published:  
— with international search report  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

[Continued on next page]

(54) Title: ADVERTISEMENT MONITORING AND FEEDBACK SYSTEM



(57) Abstract: An advertising monitoring and feedback system (100) is presented in which subscriber selections including channel changes are monitored, and in which information regarding an advertisement is extracted from text related to the advertisement. The text related to the advertisement is in the form of closed-caption text, data transmitted with the advertisement, or other associated text. A record of the effectiveness of the advertisement is created in which measurements of the percentage of the advertisement which was viewed are stored. Such records are then fed back to a subscriber characterization system (104) that incorporates advertising related information. Prior to generating subscriber profiles (106), the subscriber characterization system (104) also evaluates the subscriber selection data to eliminate data associated with irrelevant activities, such as channel surfing, channel jumping, or extended periods of inactivity.



*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**TITLE*****Advertisement Monitoring And Feedback System*****Background of the Invention**

5 Cable television service providers have typically provided one-way broadcast services but now offer high-speed data services and can combine traditional analog broadcasts with digital broadcasts and access to Internet web sites. Telephone companies can offer digital data and video programming on a  
10 switched basis over digital subscriber line technology. Although the subscriber may only be presented with one channel at a time, channel change requests are instantaneously transmitted to centralized switching equipment and the subscriber can access the programming in a broadcast-like  
15 manner. Internet Service Providers (ISPs) offer Internet access and can offer access to text, audio, and video programming which can also be delivered in a broadcast-like manner in which the subscriber selects "channels" containing programming of interest. Such channels may be offered as part  
20 of a video programming service or within a data service and can be presented within an Internet browser.

Advertisements are a part of daily life and certainly an important part of entertainment programming, where the payments for advertisements cover the cost of network television. A  
25 method, which provides a flexible billing plan to cable network users based on the amount of advertisements viewed is described in U.S. Patent No. 5,532,735, which discloses a method of advertisement selection for interactive services. A user associated with an interactive TV is presented with a program  
30 and a set of advertisements. The user can indicate the amount of advertisements in the set of advertisements he wants to view.

While advertisements are sometimes beneficial to subscribers and deliver desired information regarding specific

products or services, consumers generally view advertising as a "necessary evil" for broadcast-type entertainment. For example, a method for obtaining information on advertised services or products is described in U.S. Patent No. 5,708,478, which discloses a computer system for enabling radio listeners and television watchers to obtain advertising information. The system determines whether an incoming video or audio signal includes advertisement specific data for an advertiser and captures and stores the advertiser specific data.

Manufacturers pay an extremely high price to present, in 30 seconds or less, an advertisement for their product, which they hope a consumer will watch. Unfortunately for the manufacturer, the consumer frequently uses that interval of time to check the programming being presented on other channels, and may not watch any of the advertisement. Alternately, the consumer may mute the channel and ignore what the manufacturer has presented. In any case the probability that the consumer has watched the advertisement is quite low. It is not until millions of dollars have been spent on an advertising campaign that a manufacturer can determine that the ads have been effective. This is presently accomplished by monitoring sales of the product or TV programs or channels viewed by users as disclosed in various public documents. As an example, U.S. Patent No. 4,546,382 discloses a television and market research data collection system and method. A data collection unit containing a memory stores data as to which of the plurality of TV modes are in use, which TV channel is being viewed as well as input from a suitable optical scanning device for collecting information about the user's product purchases. Another system described in U.S. Patent No. 4,258,386 discloses a television audience measuring system. The system monitors and stores information representative of channel identification, the time at which the channel is selected and the time at which the selection of a channel is terminated.

U.S. Patent No. 5,608,445 discloses a method and device for data capture in television viewer research. Devices are attached to a video installation in order to determine to which channel a set is tuned.

5       With the advent of the Internet, manufacturers and service providers have found ways to selectively insert their advertisements based on a subscribers requests for information. As an example, an individual who searches for "cars" on the Internet may see an advertisement for a particular type of car.

10       Various internet-based advertisers use this method. The product literature from IMGIS Inc., "Ad Force," printed from the World Wide Web site [http://www.starpt.com/core/ad\\_Target.html](http://www.starpt.com/core/ad_Target.html) on June 30, 1998 discloses an ad targeting system. The system delivers ads to

15       web site visitors based on the content of the web page, time of day, day of the week, keyword, by the number of times a visitor sees an advertisement and by the order in which a series of advertisements are shown to a visitor. Nevertheless, unless the subscriber actually goes to the advertised web site, there

20       is no way to determine if the advertisement has been watched. As the content on the Internet migrates to multimedia programming including audio and video, the costs for the advertising will increase, but unless the advertiser can be sure that a significant percentage of the message was watched

25       or observed, the advertising is ineffective. Prior art products for generating reports of ad campaigns are generally PC-centric as described in various product literature which include the product literature from DoubleClick Inc., "DoubleClick: Reporting," printed from the World Wide Web (WWW)

30       site [http://www.doubleclick.net/dart/howl\\_repo.htm](http://www.doubleclick.net/dart/howl_repo.htm) on June 19, 1998, which discloses the reporting capabilities of DoubleClick's Dynamic Advertising Reporting & Targeting (DART) product. The information in the reports includes daily impressions by advertisement type, average impression per day

of week and by hour of day. The average response rate per user is also included in the reports. The product literature from Netgravity Inc. "AdServer 3," printed from the World Wide Web site <http://www.netgravity.com/products/> on July 9, 1998

- 5 discloses Netgravity's Adserver 3 product for online advertisement. The product generates reports including the profiles of visitors who viewed an ad and site traffic throughout the day, week, month and year.

- The product literature from Media Metrix "Frequently Asked Questions", printed from the World Wide Web site [http://www.mediametrix.com/interact\\_mmfaq.htm](http://www.mediametrix.com/interact_mmfaq.htm) on June 30, 1998 discloses Media Metrix's software, "PC Meter", that runs in the background of a PC and monitors everything being done on that machine. It determines who is using the PC by age, income, gender and geographic region and tracks usage of software application, commercial online services and detailed page level viewing of the World Wide Web. The marketing literature from Matchlogic Inc., "Centralized Ad Management," printed from the World Wide Web site <http://www.matchlogic.com/docs/services2.htm> on July 1, 1998 discloses Matchlogic services for ad management. The services include delivering advertisements based on pre-defined targeting criteria, generating reports on how many unique viewers saw which banner and how many times it was viewed. The product literature from Accipiter Inc., "Accipiter AdManager 2.0," printed from the World Wide Web site <http://www.accipiter.com/products/ADManager/fab.html> on July 9, 1998 discloses Accipiter's ad management system. After delivering an advertisement based on pre-defined criteria, the system can generate reports on an ad campaign. The reports include visitors' demographic data, number of impressions and clicks generated from the entire site and by each ad and advertiser.

In order to deliver more targeted programming and advertising to subscribers, it is necessary to understand their likes and dislikes to a greater extent than is presently done today. Systems which identify subscriber preferences based on their purchases and responses to questionnaires allow for the targeted marketing of literature in the mail, but do not in any sense allow for the rapid and precise delivery of programming and advertising which is known to have a high probability of acceptance to the subscriber. Other systems give users the possibility to choose their programming as described in U.S. Patent No. 5,223,924 which discloses a system and method for automatically correlating user preferences with a TV program information database. The system includes a processor that performs "free text" search techniques to correlate the downloaded TV program information with the viewer's preferences. This system requires an interaction between the users and the programming. The white paper from Net Perceptions corporation entitled "Adding Value in the Digital Age" and printed from the World Wide Web site <http://www.netperceptions.com/products/white-papers.html> on June 30, 1998 discloses how the GroupLens Recommendation Engine gives online businesses the ability to target and personalize services, content, products and advertising. A learning process learns personal information about an individual using explicit and implicit ratings, a prediction process predicts user preference using collaborative filtering and the recommendation process recommends products or services to users based on predictions.

The product literature from Aptex Software Inc., "SelectCast for Commerce Servers," printed from the World Wide Web site <http://www.aptex.com/products-selectcast-commerce.htm> on June 30, 1998 describes the product "SelectCast" for Commerce Servers. It personalizes online shopping based on observed user behavior. User interests are learned based on



the content they browse, the promotions they click and the products they purchase.

In order to determine which programming or advertising is appropriate for the subscriber, knowledge of that subscriber and the subscriber product and programming preferences is required. Different methods are being used to gain knowledge of user's preferences and to profile the users. Generally, these methods use content or data mining technologies to profile users or predict their preferences. Another technique for predicting user's preferences is based on the use of collaborative filtering as described in U.S. Patent No. 5,704,017 which discloses a collaborative filtering system utilizing a belief network. The system learns a belief network using prior knowledge obtained from an expert in a given field of decision making and a database containing empirical data such as users' attributes as well as their preferences in that decision making field. The belief network can determine the probability of the unknown preferences of the user given the known attributes and thus predicts the preference most likely to be desired by the user.

The product literature from Aptex software Inc., "SelectCast for Ad Servers," printed from the World Wide Web site <http://www.aptex.com/products-selectcast-ads.htm> on June 30, 1998 discloses an ad targeting system from Aptex Software Inc. The system employs neural networks and a context vector data model to optimize relationships between users and content. It provides user profiling by mining the context and content of all actions including clicks, queries, page views and ad impressions. Aptex's technology uses a context vector data modeling technique described in U.S. Patent No. 5,619,709 which discloses a system and method of context vector generation and retrieval. Context vectors represent conceptual relationships among information items by quantitative means. A neural network operates on a training corpus of records to develop

relationship-based context vectors based on word proximity and co-importance. Geometric relationships among context vectors are representative of conceptual relationships among their associated items.

5       The product data sheet from Open Sesame, "Learn Sesame," printed from the World Wide Web site [http://www.opensesame.com/prod\\_04.html](http://www.opensesame.com/prod_04.html) on July 09, 1998 discloses Open Sesame's personalization product for Web enterprises. It learns about users automatically from their  
10       browsing behavior.

      The product literature from Engage Technologies, "Engage.Discover," printed from the World Wide Web site <http://www.engagetech.com> on July 09, 1998 discloses Engage Technologies' product for user profiling. User-disclosed  
15       information such as interest, demographics and opinions are combined with anonymous clickstream data that describes where users come from before visiting the site, how long they stay, and what pages or types of pages they visit most frequently to build the visitor profile.

20       The marketing literature from BroadVision, "The Power of Personalization", printed from the World Wide Web site <http://www.broadvision.com/content/corporate/brochure/Broch4.htm> on August 21, 1998 discloses the BroadVision One-to-One application profiling system. The system learns about users  
25       through a variety of techniques including registration, questionnaires, observation and integration of historical and externally generated data.

      The marketing literature from the Firefly Corporation, "Firefly Passport Office," printed from the World Wide Web site  
30       <http://www.firefly.net/company/PassportOffice.html> on June 20, 1998 discloses Firefly's Relationship Management software. The software enables online businesses to create, extend and manage personal profiles for every user.

Specific information regarding a subscriber's viewing habits or the Internet web sites they have accessed can be stored for analysis, but such records are considered private and subscribers are not generally willing to have such  
5 information leave their control. Although there are regulatory models, which permit the collection of such data on a "notice and consent" basis, there is a general tendency towards legal rules, which prohibit such raw data from being collected.

With the migration of services from a broadcast based  
10 model to a client-server based model in which subscribers make individualized requests for programming to an Internet access provider or content provider, there is opportunity to monitor the subscriber viewing characteristics to better provide them with programming and advertising which will be of interest to  
15 them. A server may act as a proxy for the subscriber requests and thus be able to monitor what a subscriber has requested and is viewing. Since subscribers may not want this raw data to be utilized, there is a need for a system which can process this information and generate statistically relevant subscriber  
20 profiles. These profiles should be accessible to others on the network who may wish to determine if their programming or advertisements are suitable for the subscriber. In a broadcast-based model, the information to be processed can be embedded within the TV program or broadcast separately and can  
25 be in the form of an electronic program guide (EPG) or text information related to the program. As an example, U.S. Patent No. 5,579,055 discloses an electronic program guide (EPG) and a text channel data controller. The text and EPG data are embedded in the vertical blanking interval of the video signal  
30 and extracted, at reception, by the data controller. The EPG contains information fields such as program category, program subcategory and program content description. U.S. Patent No. 5,596,373 discloses also a method and apparatus for providing program oriented information in a multiple station broadcasting

system. The EPG data includes guide data, channel data and program data. The program data includes among other information, the program title, the program category, the program sub-category and a detailed description of the program.

5       Some efforts have been made to transmit targeted advertisements and thereupon collect feedback on the transmitted advertisements. For example, U.S. Patent No. 5,948,061 provides methods and apparatuses for targeting the delivery of advertisements over a network such as the Internet.  
10       Statistics are compiled on individual users and networks and the use of the advertisements is tracked to permit targeting of the advertisements to individual users. In response to requests from affiliated sites, an advertising server transmits to people accessing the page of a site an appropriate one of  
15       the advertisements based upon profiling of users and networks.

      Furthermore, U.S. Patent No. 6,005,597 provides a method and apparatus for television program selection that monitors the viewing preferences of a viewer to create a dynamic viewer profile that is used to rate available programs. Based on the  
20       viewer profile, available programs are sorted and presented to the viewer in descending order of predicted interest. The invention allows a viewer to quickly find the program of greatest interest to the viewer without having to tediously search through large numbers of available programs. The  
25       invention may also be used for selecting from among a plurality of programs other than television programs, such as, for example, radio programs or audio or video programs stored on digital storage media such as CD's and DVD's.

      U.S. Patent No. 5,446,919 provides a communication system  
30       capable of targeting a demographically or psychographically defined audience. A master database is maintained, containing demographic and psychographic information about each audience member. This information is transmitted and stored in a channel selection/decoder unit associated with each audience

member's receiver. Multiple media messages are transmitted to each audience member. Accompanying the transmission is a selection profile command, which details the demographic/psychographic profile of audience members that are to receive each media message. The channel selector/decoder unit associated with each member's receiver compares the selection profile with the demographic/psychographic information about the audience member and selects the appropriate media message for that audience member.

U.S. Patent Nos. 5,991,735 and 5,848,396 disclose a computer network method and apparatus which provides targeting of an appropriate audience based on psychographic or behavioral profiles of end users. The psychographic profile is formed by recording computer activity and viewing habits of the end user. Content of categories of interest and display format in each category are revealed by the psychographic profile, based on user viewing of aggregate information. Using the profile (with or without additional user demographics), advertisements are displayed to appropriately selected users. Based on regression analysis of recorded responses of a first set of users viewing the advertisements, the target user profile is refined. Viewing by and regression analysis of recorded responses of subsequent sets of users continually auto-targets and customizes ads for the optimal end user audience.

U.S. Patent No. 5,155,591 provides a functionality where different commercial messages are broadcast to different demographically targeted audiences in a cable television system or the like. A first television channel contains television programs and periodic commercial messages. A second television channel contains alternate commercial messages. Demographic characteristics of a viewer are identified, and commercial messages are selectively provided from the first or second channel, depending upon the viewer's demographic characteristics. Demographic data can be input by a viewer via

a remote control, downloaded to a subscriber's converter from a remote head-end, or programmed into the converter at installation. Prioritization of the demographic characteristics of a plurality of television viewers watching a program together enables commercials to be targeted to the viewer having highest priority. Statistical data can be maintained concerning the number and identity of subscribers viewing specific commercials.

U.S. Patent No. 5,774,170 enhances television (and radio) advertising by targeting, delivering and displaying electronic advertising messages (commercials) within specified programming in one or more pre-determined households (or on specific display devices) while simultaneously preventing a commercial from being displayed in other households or on other displays for which it is not intended. Commercials can be delivered to specified homes or displays via either over-the-air or wired delivery systems.

U.S. Patent No. 5,446,919 provides a communication system capable of targeting a demographically or psychographically defined audience. A master database is maintained, containing demographic and psychographic information about each audience member. This information is transmitted and stored in a channel selection/decoder unit associated with each audience member's receiver. Multiple media messages are transmitted to each audience member. Accompanying the transmission is a selection profile command, which details the demographic/psychographic profile of audience members that are to receive each media message. The channel selector/decoder unit associated with each member's receiver compares the selection profile with the demographic/psychographic information about the audience member and selects the appropriate media message for that audience member.

IEEE Publication: "A Framework for Targeting Banner Advertising on the Internet" by K.Gallagher and J. Parsons,

Jan. 1997, presents a model for effectively and efficiently targeting hypermedia-based banner advertisements in an online information service. The model takes advantage of information technology to micro-target banner advertisements based on  
5 individual characteristics of users.

U.S. Patent No. 5,636,346 provides a method and system for selectively targeting advertisements and programming to different demographically and specially targeted television audiences by relating carrier subscriber data to other  
10 proprietary marketing databases by creating, compiling and updating a National Directory of cable and other carrier system subscriber names and address information derived from actual cable system and telephone company billing records.

The above-mentioned systems and methods provide many  
15 different ways to select and deliver targeted advertisements. However, these systems lack the means for monitoring displayed advertisements and collecting feedback. Such monitoring and feedback systems are typically not available in television environments.

For the foregoing reasons, there is a need for an  
20 advertisement monitoring and feedback system that may monitor advertisements that have been viewed by a subscriber and provide feedback to a subscriber characterization system which may generate one or more subscriber profiles that incorporate  
25 advertising related information. Such subscriber profiles may assist advertisers in determining the success rates of the advertisements being broadcast and accordingly advertisers may allocate financial sources on the advertisements that are likely to succeed in getting the attention of the subscribers.

30

#### Summary of the Invention

The present invention encompasses an advertisement monitoring system that determines to what extent an advertisement has been viewed by a subscriber or a household,

and generates one or more records of the subscriber advertising selection data. In a preferred embodiment, the subscriber advertising selection data includes the channel selected and the time at which the channel was selected. Advertisement  
5 related information including the type of product, brand name, and other descriptive information that categorizes the advertisement is also extracted and added to the advertising selection data. The advertisement related information might also be extracted from the closed-captioning text. Based on  
10 this advertisement selection data, a record is created documenting what percentage, if any, of the advertisement was watched. This record is then fed back to the subscriber characterization system, wherein the subscriber characterization system may comprise one or more filters.

15 The subscriber characterization system also monitors the subscriber viewing habits associated with program viewing and generates one or more records of program selection choices. The program selection choices include the viewing time duration, number of channel changes, volume at which the  
20 programming is listened, program selection, and text information about the programming. The records are used to determine what type of programming the subscriber is most interested in. The subscriber characterization system characterizes subscribers based on feedback information from an  
25 advertisement monitoring system and monitors their detailed program viewing selections.

Furthermore, the subscriber characterization system is equipped with one or more filters that assist in determining and eliminating from consideration, selection data associated  
30 with irrelevant activities by the subscriber. Examples of irrelevant activities include selection data associated with channel surfing and/or channel jumping (up and down) activities.

The channel surfing activity refers to one or more rapid



channel changes initiated by the subscriber for the purpose of selecting a channel/program for actual viewing. Generally, the subscriber selects a channel, and views the contents of the program at the selected channel for a few seconds (about 3-4  
5 seconds), and then changes the channel to view the contents of the next channel. Such rapid changes generally occur a few times in a row before the subscriber selects a channel/programming for actual viewing. The filters of the present invention are configured to detect channel surfing  
10 activities by the subscriber by monitoring and evaluating associated viewing times. The channel surfing activities are filtered out and thus not considered in the determination of actual viewing selections.

The channel jumping refers to an activity wherein the  
15 subscriber changes channels very rapidly in order to move from an existing channel to a desired channel. Therein, the subscriber is not channel surfing, instead the subscriber already knows the intended channel/program for actual viewing and is jumping channels to reach the desired channels. For  
20 example, if the subscriber is at channel number 6, and wants to go to channel number 12, the subscriber may jump the channel by changing the channel six times. Generally, in channel jumping, the channel changes occur very rapidly and the viewing time at the each channel is very brief, e.g., less than one second.  
25 The filters of the present invention are configured to detect and filter out channel jumping. Thus, the channel jumping activities are not considered in the determination of actual viewing selections.

The filters of the present invention are also capable of  
30 monitoring extended spans of inactivity, e.g., a lack of any channel changes, volume changes, or any other selection changes activity for more than 3 hours. Such spans of inactivity are considered "dead periods" implying that the subscriber is not actively watching the video and/or other multimedia

programming. The reasons for such dead periods may be caused by the fact that the subscriber has left the room, or the subscriber is not active (e.g., the subscriber has gone to sleep or has dozed off), or the fact that the subscriber is  
5 actively engaging in another activity within the room and is not attending to the programming.

The subscriber characterization system of the present invention analyzes the actual viewing selections made by the subscriber or the subscriber household, and generates a  
10 demographic description of the subscriber or household. This demographic description describes the probable age, income, gender and other demographics. The resulting characterization includes probabilistic determinations of what other programming or products in which the subscriber/household will be  
15 interested.

In a preferred embodiment, the text information related to the advertisement is processed using context mining techniques which allow for classification of the advertisement and extraction of key data including product type and brand.  
20 Context mining techniques allow for determination of a product type, product brand name and in the case of a product which is not sold with a particular brand name, a generic name for the product.

The present invention can also be realized in a client-server mode in which case the advertising related information is collected at the client side of the network and is then transmitted to the server side via a secure connection. The server side then incorporates this information with other subscriber information to create subscriber profiles.  
25

30 These and other features and objects of the invention will be more fully understood from the following detailed description of the preferred embodiments which should be read in light of the accompanying drawings.

### Brief Description of the Drawings

The accompanying drawings, which are incorporated in and form a part of the specification, illustrate the embodiments of the present invention and, together with the description serve to explain the principles of the invention.

In the drawings:

FIG. 1A illustrates a context diagram for an advertising monitoring and feedback system (AMFS), in accordance with one embodiment of the present invention;

10 FIG. 1B illustrates a context diagram for an advertising monitoring system (AMS), in accordance with one embodiment of the present invention;

FIG. 1C illustrates a context diagram for a subscriber characterization system with filters (SCSF), in accordance with one embodiment of the present invention;

FIG. 1D illustrates a functional diagram of the processing utilized by the filters;

FIG. 2 illustrates a block diagram for a realization of a subscriber monitoring system for receiving video signals;

20 FIG. 3 illustrates a block diagram of a channel processor;

FIG. 4 illustrates a block diagram of a computer for a realization of the advertisement monitoring system;

FIG. 5 illustrates a channel sequence and volume over a twenty-four (24) hour period;

25 FIG. 6A illustrates a time of day detailed record;

FIG. 6B illustrates the processing utilized to determine channel surfing activities;

FIG. 6C illustrates the processing utilized by filters to determine channel-jumping activities;

30 FIG. 7 illustrates a household viewing habits statistical table;

FIG. 8A illustrates an entity-relationship diagram for the generation of program characteristics vectors;

FIG. 8B illustrates a flowchart for program characterization;

FIGS. 9A illustrates a deterministic program category vector;

5       FIG. 9B illustrates a deterministic program sub-category vector;

FIG. 9C illustrates a deterministic program rating vector;

FIG. 9D illustrates a probabilistic program category vector;

10       FIG. 9E illustrates a probabilistic program sub-category vector;

FIG. 9F illustrates a probabilistic program content vector;

FIG. 10A illustrates a set of logical heuristic rules;

15       FIG. 10B illustrates a set of heuristic rules expressed in terms of conditional probabilities;

FIG. 11 illustrates an entity-relationship diagram for the generation of program demographic vectors;

FIG. 12 illustrates a program demographic vector;

20       FIG. 13 illustrates an entity-relationship diagram for the generation of household session demographic data and household session interest profiles;

FIG. 14 illustrates an entity-relationship diagram for the generation of average and session household demographic characteristics;

FIG. 15 illustrates average and session household demographic data;

FIG. 16 illustrates an entity-relationship diagram for generation of a household interest profile;

30       FIG. 17 illustrates a household interest profile including programming and product profiles;

FIG. 18 illustrates a client-server architecture for realizing the present invention; and

FIG. 19 illustrates an advertisement monitoring table.

**Detailed Description  
Of The Preferred Embodiment**

In describing a preferred embodiment of the invention  
5 illustrated in the drawings, specific terminology will be used  
for the sake of clarity. However, the invention is not  
intended to be limited to the specific terms so selected, and  
it is to be understood that each specific term includes all  
technical equivalents which operate in a similar manner to  
10 accomplish a similar purpose.

With reference to the drawings, in general, and FIGS. 1  
through 19 in particular, the apparatus of the present  
invention is disclosed.

FIG. 1A illustrates a context diagram of a preferred  
15 embodiment of an Advertising Monitoring and Feedback System  
(AMFS) 100. The AMFS 100 comprises an advertising monitoring  
system (AMS) 102 and a subscriber characterization system with  
filters (SCSF) 104. A context diagram, in combination with  
entity-relationship diagrams, provide a basis from which one  
20 skilled in the art can realize the present invention. The  
present invention can be realized in a number of programming  
languages including C, C++, Perl, and Java, although the scope  
of the invention is not limited by the choice of a particular  
programming language or tool. Object oriented languages have  
25 several advantages in terms of construction of the software  
used to realize the present invention, although the present  
invention can be realized in procedural or other types of  
programming languages known to those skilled in the art.

Within the AMFS 100, the AMS 102 monitors subscriber  
30 advertising viewing activities, processes them, and generates  
advertising selection data. The AMS 102 then forwards the  
advertising selection data to the SCSF 104 which further  
processes the advertising related information received from the  
AMS 102. The SCSF 104 also monitors subscriber viewing habits

as they relate to actual programming information. SCSF 104, based on advertising related information and the actual programming information, generates one or more subscriber profiles 106, wherein each subscriber profile indicates a  
5 probabilistic measure of subscriber demographics and/or preferences.

As illustrated in FIG. 1B, in generating one or more records of advertisement selections, the AMS 102 receives from a subscriber 120A commands in the form of a volume control  
10 signal 124A or advertising selection data 122A which can be in the form of a channel change. The advertising material being viewed by the subscriber 120A is referred to as source material 130A. The source material 130A, as defined herein, is the content that a subscriber selects and may consist of analog  
15 video, Motion Picture Expert Group (MPEG) digital video source material, other digital or analog material, Hypertext Markup Language (HTML) or other type of multimedia source material. The AMS 102 can access the source material 130A received by the subscriber 120A using a start signal 132A and a stop signal  
20 134A, which control the transfer of source related text 136A which can be analyzed as described herein.

In a preferred embodiment, the source related text 136A can be extracted from the source material 130A and stored in memory. The source related text 136A, as defined herein,  
25 includes source related textual information including descriptive fields which are related to the source material 130A, or text which is part of the source material 130A itself. The source related text 136A can be derived from a number of sources including but not limited to closed-captioning  
30 information, Electronic Program Guide (EPG) material, and text information in the source itself (e.g. text in HTML files).

An Electronic Program Guide (EPG) 140A contains information related to the source material 130A which is useful to the subscriber 120A. The EPG 140A is typically a

navigational tool which contains source related information, including but not limited to, the programming category, program description, rating, actors, and duration. The structure and content of EPG data is described in detail in US Patent

5 5,596,373 assigned to Sony Corporation and Sony Electronics. The EPG 140A can be accessed by the AMS 102 by a request EPG data signal 142A which results in the return of a category 144A, a sub-category 146A, and a general advertisement (ad) description 148A. The EPG 140A can potentially include fields  
10 related to advertising. In one embodiment of the present invention, the EPG information such as the category 144A, the sub-category 146A, and the general advertisement description 148A are stored in a memory.

In another embodiment of the present invention, the source  
15 related text 136A is the closed-captioning text embedded in the analog or digital video signal. Such closed-captioning text can be stored in a memory for processing to extract the advertising characteristics related information.

One of the functions of the AMS 102 is to generate one or  
20 more records of advertisement selections 110A which are comprised of advertising characteristics. The advertising selection record 110A may be obtained from the monitored activities of the subscriber 120A and in a preferred embodiment can be stored in a dedicated memory. In an alternate  
25 embodiment, the advertising selection record 110A may be stored in a storage disk. In an exemplary case, the information which is utilized to form the advertisement selection data 110A includes time 112A, which corresponds to the time of an event, channel ID 114A, program ID 116A, volume level 118A, channel  
30 change record 119A, and advertising title 117A. The AMS 102 forwards the advertising selection records 110A to the SCSF 104 for further processing and for generating subscriber profiles.

FIG. 1C depicts the context diagram of a preferred embodiment of the SCSF 104.

In the process of collecting raw subscriber selection data, the SCSF 104 receives from a subscriber 120B commands in the form of a volume control signal 124B or program selection data 122B which can be in the form of a channel change but may  
5 also be an address request which requests the delivery of programming from a network address. A record signal 126B indicates that the programming or the address of the programming is being recorded by the subscriber 120B. The record signal 126B can also be a printing command, a tape  
10 recording command, a bookmark command or any other command intended to store the program being viewed, or program address, for later use.

The material being viewed by the subscriber 120B is referred to as source material 130B. The source material 130B,  
15 as defined herein, is the content that a subscriber selects and may consist of analog video, digital video such as MPEG digital video, Hypertext Markup Language (HTML), or other types of multimedia source material. The SCSF 104 can access the source material 130B received by the subscriber 120B using a start  
20 signal 132B and a stop signal 134B, which control the transfer of source related text 136B which can be analyzed as described herein.

In a preferred embodiment, the source related text 136B can be extracted from the source material 130B and stored in  
25 memory. The source related text 136B, as defined herein, includes source related textual information including descriptive fields which are related to the source material 130B, or text which is part of the source material 130B itself. The source related text 136B can be derived from a number of  
30 sources including but not limited to closed-captioning information, EPG material, and text information in the source itself (e.g., text in HTML files).

The EPG 140B contains information related to the source material 130B which is useful to the subscriber 120B. The EPG



140B is typically a navigational tool which contains source related information including but not limited to the programming category, program description, rating, actors, and duration. The EPG 140B can be accessed by the SCSF 104 by a  
5 request EPG data signal 142B which results in the return of a category 144B, a sub-category 146B, and a program description 148B. The EPG data may be stored in memory.

In another embodiment of the present invention, the source related text 136B is the closed-captioning text embedded in the  
10 analog or digital video signal. Such closed-captioning text can be stored in memory for processing to extract characteristic vectors 150.

The SCSF 104 generates one or more program selection records 110B by monitoring and storing activities of the  
15 subscriber 120B. In an exemplary case, the program selection records 110B include time 112B, which corresponds to the time of an event, channel ID 114B, program ID 116B, program title 117B, volume level 118B, and channel change record 119B. A detailed record of selection data is illustrated later.  
20 Generally, the program selection records 110B contain the raw data accumulated over a predetermined period of time and relate to viewing selections made by the subscriber 120B over the predetermined period of time.

The system of the present invention also comprises one or  
25 more filters 115A, 115B. The filters 115A, 115B evaluate the advertisement and program selection data and eliminate any selection data associated with irrelevant activities, and in turn generate actual subscriber selection data 199 that corresponds only to the actual viewing selections made by the  
30 subscriber 120B. The filters 115A, 115B may be a computer means or a software module configured with some predetermined rules. These predetermined rules assist in recognizing irrelevant activities and the elimination of the selection data from the raw subscriber selection data. In one embodiment, the

selection data associated with channel surfing, channel jumping and dead periods is eliminated from the advertisement and program selection data to generate actual subscriber selection data 199. FIG. 1D and the related text describe this process  
5 in detail.

The actual subscriber selection data 199 comprises time 112C, which corresponds to the time of an actual viewing event exclusive of channel surfing, channel jumping or dead periods, channel ID 114C, program ID 116C, program title 117C, volume  
10 level 118C, and channel change record 119C.

Based on the actual subscriber selection data 199, the SCSF 104 generates one or more program characteristics vectors 150 which are comprised of collected characteristics data 152. The characteristics data 152, which can be used to create the  
15 program characteristics vectors 150, both in vector and table form, are examples of source related information which represent characteristics of the source material (for the advertisements and the programs). The characteristics vectors 150 are derived from the source related text 136B and/or from  
20 the EPG 140B by applying information retrieval techniques. The details of this process are discussed in detail later.

In a preferred embodiment, the characteristics vectors 150 are lists of values which characterize the programming (source) material in accordance to the category 144B, the sub-category  
25 146B, and the program description 148B. The present invention may also be applied to advertisements, in which case program characteristics vectors contain, as an example, a product category, a product sub-category, and a brand name.

As illustrated in FIG. 1C, the SCSF 104 uses heuristic  
30 rules 160. The heuristic rules 160, as described herein, are composed of both logical heuristic rules as well as heuristic rules expressed in terms of conditional probabilities. The heuristic rules 160 can be accessed by the SCSF 104 via a request rules signal 162 which results in the transfer of a

copy of rules 164 to the SCSF 104.

The SCSF 104 also forms demographic vectors 170 from demographics data 172. The program demographic vectors 170 also represent characteristics of source related information in the form of the intended or expected demographics of the audience for which the source material is intended. The characteristics vector 150 is used in combination with the set of the heuristic rules 160 to define the demographic vectors 170.

10 In a preferred embodiment, household viewing data 197 is computed from the actual subscriber selection data 199. The household viewing data 197 is derived from the actual subscriber selection data 199 by looking at viewing habits at a particular time of day over an extended period of time, usually  
15 several days or weeks, and making some generalizations regarding the viewing habits during that time period. The SCSF 104 also transforms household viewing data 197 to form household viewing habits 195, i.e. statistical representation of subscriber/household viewing data illustrating patterns in  
20 viewing.

The SCSF 104 generates household profiles including a household demographic characteristics 190 and a household interest profile 180. The household demographic characteristics 190 resulting from the transfer of household  
25 demographic data 192, and the household interest profile 180, results from the transfer of household interests data 182. Both the household demographics characteristics 190 and the household interest profile 180 have a session value and an average value, as will be discussed herein.

30 Referring now to FIG. 1D, exemplary processing of the filters 115 is shown. As mentioned before, filters 115A and filter 115B evaluate the advertisement selection data 110A and the program selection data 110B to determine any data associated with irrelevant selection activities and then

generate actual subscriber selection data 199 which does not include irrelevant selection data. The irrelevant selection data generally corresponds to channel surfing, channel jumping, or dead periods activities. These activities are generally  
5 recognized by reviewing corresponding viewing times. In the case of channel surfing or channel jumping, the associated viewing times are very brief, a few milliseconds or a few seconds. In the case of dead periods, the viewing time is relatively long having no actions, e.g., a few hours.

10 FIG. 2 illustrates an exemplary system for monitoring subscriber activities, including advertising viewing habits, and can be used to realize the AMS 102 and the SCSF 104. In a preferred embodiment, the monitoring system of FIG. 2 is located in a television set-top device or in the television  
15 itself. In an alternate embodiment, the monitoring system is part of a computer which receives programming from a network.

In an application of the system for television services, an input connector 220 accepts the video signal coming either from an antenna, cable television input, or other network. The  
20 video signal can be analog or digital, such as MPEG. Alternatively, the video source may be a video stream or other multimedia stream from a communications network including the Internet.

In the case of either analog or digital video, selected  
25 fields are defined to carry EPG data or closed-captioning text. For analog video, the closed-captioning text is embedded in the vertical blanking interval (VBI). As described in US Patent 5,579,005, assigned to Scientific-Atlanta, Inc., the EPG information can be carried in a dedicated channel or embedded  
30 in the VBI. For digital video, the closed-captioning text is carried as video subscriber bits in a subscriber\_data field. The EPG data is transmitted as ancillary data and is multiplexed at the transport layer with the audio and video data.

In FIG. 2, a system control unit 200 receives commands from the subscriber 120, decodes the command and forwards the command to the destined module. In a preferred embodiment, the commands are entered via a remote control to a remote receiver 5 205 or a set of selection buttons 207 available at the front panel of the system control unit 200. In an alternate embodiment, the subscriber 120 enters the commands via a keyboard.

The system control unit 200 also contains a Central 10 Processing Unit (CPU) 203 for processing and supervising all of the operations of the system control unit 200, a Read Only Memory (ROM) 202 containing the software and fixed data, a Random Access Memory (RAM) 204 for storing data. The CPU 203, the RAM 204, the ROM 202, and an input/output (I/O) controller 15 201 are attached to a master bus 206. A power supply in the form of a battery can also be included in the system control unit 200 for backup in case of power outage.

The I/O controller 201 interfaces the system control unit 200 with external devices. In a preferred embodiment, the I/O 20 controller 201 interfaces to the remote receiver 205 and a selection button such as the channel change button on a remote control. In an alternate embodiment, it can accept input from a keyboard or a mouse.

The program selection data is forwarded to a channel 25 processor 210. The channel processor 210 tunes to a selected channel and the media stream is decomposed into its basic components: the video stream, the audio stream, and the data stream. The video stream is directed to a video processor module 230 where it is decoded and further processed for 30 display to the TV screen. The audio stream is directed to an audio processor 240 for decoding and output to the speakers.

The data stream can be EPG data, closed-captioning text, Extended Data Service (EDS) information, a combination of these, or an alternate type of data. In the case of EDS the

call sign, program name and other useful data are provided. In a preferred embodiment, the data stream is stored in a reserved location of the RAM 204. In an alternate embodiment, a magnetic disk is used for data storage. The system control unit 200 writes in a dedicated memory, which in a preferred embodiment is the RAM 204, the selected channel, the time of selection, the volume level and the program ID and the program title. Upon receiving the program selection data, the new selected channel is directed to the channel processor 210 and the system control unit 200 writes to the dedicated memory the channel selection end time and the program title at the time of channel change. The system control unit 200 keeps track of the number of channel changes occurring during the viewing time via the channel change record. This data forms part of the advertising selection records 110A or the programming selection records 110B.

The volume control signal is sent to the audio processor 240. In a preferred embodiment, the volume level selected by the subscriber corresponds to the listening volume. In an alternate embodiment, the volume level selected by the subscriber represents a volume level to another piece of equipment such as an audio system (home theatre system) or to the television itself. In such a case, the volume can be measured directly by a microphone or other audio sensing device which can monitor the volume at which the selected source material is being listened.

A program change occurring while watching a selected channel is also logged by the system control unit 200. Monitoring the content of the program at the time of the program change can be done by reading the content of the EDS. The EDS contains information such as program title, which is transmitted via the VBI. A change on the program title field is detected by the monitoring system and logged as an event. In an alternate embodiment, an EPG is present and program

information can be extracted from the EPG. In a preferred embodiment, the programming data received from the EDS or EPG permits distinguishing between entertainment programming and advertisements.

5        FIG. 3 illustrates an exemplary block diagram of the channel processor 210. In a preferred embodiment, the input connector 220 connects to a tuner 300 which tunes to the selected channel. A local oscillator can be used to heterodyne the signal to the IF signal. A demodulator 302 demodulates the  
10        received signal and the output is fed to an FEC decoder 304. The data stream received from the FEC decoder 304 is, in a preferred embodiment, in an MPEG format. In a preferred embodiment, a system demultiplexer 306 separates out video and audio information for subsequent decompression and processing,  
15        as well as ancillary data which can contain program related information.

      The data stream presented to the system demultiplexer 306 consists of packets of data including video, audio and ancillary data. The system demultiplexer 306 identifies each  
20        packet from the stream ID and directs the stream to the corresponding processor. The video data is directed to the video processor module 230 and the audio data is directed to the audio processor 240. The ancillary data can contain closed-captioning text, emergency messages, program guide, or  
25        other useful information.

      Closed-captioning text is considered to be ancillary data and is thus contained in the video stream. The system demultiplexer 306 accesses the subscriber data field of the video stream to extract the closed-captioning text. The  
30        program guide, if present, is carried on data stream identified by a specific transport program identifier.

      In an alternate embodiment, analog video may be used. For analog programming, ancillary data such as closed-captioning text or EDS data are carried in a vertical blanking interval.

FIG. 4 illustrates an exemplary block diagram of a computer system for a realization of the subscriber monitoring system illustrated in FIG. 2. A system bus 422 transports data amongst the CPU 203, the RAM 204, a Read Only Memory - Basic Input Output System (ROM-BIOS) 406 and other components. The CPU 203 accesses a hard drive 400 through a disk controller 402. The standard input/output devices are connected to the system bus 422 through the I/O controller 201. A keyboard is attached to the I/O controller 201 through a keyboard port 416 and the monitor is connected through a monitor port 418. The serial port device uses a serial port 420 to communicate with the I/O controller 201. Industry Standard Architecture (ISA) expansion slots 408 and Peripheral Component Interconnect (PCI) expansion slots 410 allow additional cards to be placed into the computer. In a preferred embodiment, a network card is available to interface a local area, wide area, or other network.

FIG. 5 illustrates a channel sequence and volume over a twenty-four (24) hour period associated with advertisement selection record 110A or subscriber selection record 110B. The Y-axis represents the status of the receiver in terms of on/off status and volume level. The X-axis represents the time of day. The channels viewed are represented by the windows 501-506, with a first channel 502 being watched followed by the viewing of a second channel 504, and a third channel 506 in the morning. In the evening a fourth channel 501 is watched, a fifth channel 503, and a sixth channel 505. A channel change is illustrated by a momentary transition to the "off" status and a volume change is represented by a change of level on the Y-axis.

A detailed record of the advertising selection record 110A and/or programming selection data 110B is illustrated in FIG. 6A in a table format. A time column 602 contains the starting time of every event occurring during the viewing time. A



Channel ID column 604 lists the channels viewed or visited during that period. A title column 603 contains the titles of programs/advertisements viewed. A volume column 601 contains the volume level at the time of viewing a selected channel.

5        Generally, the advertisement selection record 110A and/or program selection record 110B are unprocessed data and comprises the data associated with irrelevant or inconsequential activities, e.g., channel surfing, channel jumping, or dead activities. Thus, before the subscriber/  
10 household viewing habits are determined, the raw selection data is filtered to eliminate the data associated with irrelevant (inconsequential) activities such as channel surfing, channel jumping, or dead period activities. It is to be noted that raw selection data refers back to advertisement selection records  
15 110A and program selection records 110B.

As illustrated in FIG. 6B, the channel surfing relates to an activity wherein the subscriber rapidly changes channels before arriving at a channel which may be of interest to him. During the channel surfing period, the viewing time of each  
20 intermediate channel is very brief, e.g., less than one minute. In this viewing time, the subscriber briefly glances at the channel programming, and then moves on to the next channel.

One or more filters of the present invention are configured to filter out the surfing activity and only the  
25 actual viewing activity is considered in the actual make-up of household viewing habits. For example, in FIG. 6B, the viewing record illustrates that the viewing time of each of the channels 2, 3, 4, 5 is less than a minute, however, the viewing time of channel 6 is about an hour. Filter 115 of the present  
30 invention evaluates this record, and then removes the corresponding viewing times of channel 2, 3, 4, 5 from the viewing records. The viewing time of channel number 6 is kept as it is not indicative of the channel surfing, but an actual viewing.

Similarly, the viewing record also indicates that the, corresponding viewing times of each of channel numbers 7, 8, 9, 58, 57, 56, 55, 54, 53 are about a minute or less. This implies that after the subscriber had completed the viewing of  
5 channel number 6, the subscriber once again surfed the channels to find a programming of interest at channel 25, which was viewed for about 10 minutes.

FIG. 6C illustrates processing involved in the elimination of viewing times associated with the channel jumping  
10 activities. The channel jumping activity is different than a channel surfing activity in a sense that the subscriber already knows the intended programming (and corresponding channel number) he wants to watch, and utilizes the channel up or channel down button to arrive at the intended channel.

15 The viewing time of all the intermediate channels during channel jumping activity are generally very brief (less than a second). Also, as the channel up or channel down button is utilized to reach the desired channels, generally, there exists an upwards or a downwards stream of channel changes, i.e., the  
20 subscriber may jump through channels 2, 3, 4 and 5 to reach channel number 6 (an intended channel). Similarly, subscriber jumps may through channel 7, 8, 9, 10, 11, 12, 13, 14, 15, and 16 to reach channel 17.

The filters of the present invention are configured to  
25 eliminate the channel jumping data from the actual viewing data. Filters generally evaluate the associated viewing times, and all the viewing times which correspond to channel jumping, e.g., are less than one second, are removed from the viewing records.

30 The filters are also configured to eliminate data associated with dead activities, e.g., extended spans of inactivity. These extended spans of inactivity indicate that the subscriber is not actively watching the programming, e.g., the subscriber has left the room, has gone to sleep, or is

otherwise engaged in some other activity. These spans of inactivity may be determined by evaluating channel change commands, volume change commands, or other program selection commands issued by the subscriber. For example, if the  
5 evaluation of the viewing record indicates that the subscriber has not issued either of the channel change, volume change, on/off, or any other program selection command in last three hours, it is assumed that subscriber is in an inactive condition, and the remaining viewing time of that viewing  
10 session is not considered in the make-up of the household viewing habits. The spans of inactivity may be caused by any number of reasons. For example, it is generally known that subscribers often do not turn their televisions (or other multimedia sources) off before attending to some other  
15 activities, e.g. cooking in the kitchen, running to the nearby grocery store, or going to basement for a work-out.

The filters of the present invention are constantly filtering out the irrelevant information associated with the channel surfing activities, channel jumping activities, or with  
20 the periods of inactivity, so that the data used for generating household viewing habits is more illustrative of the actual viewing habits. The actual subscriber selection data is then used to create household viewing habits.

A representative statistical record corresponding to the  
25 household viewing habits is illustrated in FIG. 7. In a preferred embodiment, a time of day column 700 is organized in period of time including morning, mid-day, afternoon, night, and late night. In an alternate embodiment, smaller time periods are used. A minutes watched column 702 lists, for each  
30 period of time, the time in minutes in which the SCSF 104 recorded delivery of programming. The number of channel changes during that period and the average volume are also included in that table in a channel changes column 704 and an average volume column 706 respectively. The last row of the

statistical record contains the totals for the items listed in the minutes watched column 702, the channel changes column 704 and the average volume 706.

FIG. 8A illustrates an entity-relationship diagram for the generation of the characteristics vector 150. The context vector generation and retrieval technique can be applied for the generation of the characteristics vectors 150 which is described in US Patent 5,619,709. Other techniques are well known by those skilled in the art.

The source material 130 or the EPG 140 is passed through a characterization processing 800 to generate the characteristics vectors 150. The characterization processing 800 is described in accordance with FIG. 8B. The content descriptors including a first content descriptor 802, a second content descriptor 804 and an nth content descriptor 806, each classified in terms of the category, the sub-category, and other divisions as identified in the industry accepted program classification system, are presented to a context vector generator 820. As an example, the content descriptor can be text representative of the expected content of material found in the particular category (program and/or advertisement). In this example, the content descriptors 802, 804 and 806 would contain text representative of what would be found in programs in the news, fiction, and advertising categories respectively. The context vector generator 820 generates context vectors for that set of sample texts resulting in a first summary context vector 808, a second summary context vector 810, and an nth summary context vector 812. In the example given, the summary context vectors 808, 810, and 812 correspond to the categories of news, fiction and advertising respectively. The summary vectors are stored in a local data storage system.

A sample of the source related text 136 which is associated with the new program to be classified is passed to the context vector generator 820 which generates a context

vector 840 for that program. The source related text can be either the source material 130, the EPG 140, or other text associated with the source material 130. A comparison is made between the actual program context vectors and the stored  
5 program content context vectors by computing, in a dot product computation process 830, the dot product of the first summary context vector 808 with the context vector 840 to produce a first dot product 814. Similar operations are performed to produce second dot product 816 and nth dot product 818.

10 The values contained in the dot products 814, 816 and 818, while not probabilistic in nature, can be expressed in probabilistic terms using a simple transformation in which the result represents a confidence level of assigning the corresponding content to that program. The transformed values  
15 add up to one. The dot products can be used to classify a program, or form a weighted sum of classifications which results in the characteristics vectors. In the example given, if the source related text 136 was from an advertisement, the nth dot product 818 would have a high value, indicating that  
20 the advertising category was the most appropriate category, and assigning a high probability value to that category. If the dot products corresponding to the other categories were significantly higher than zero, those categories would be assigned a value, with the result being a program  
25 characteristics vector 150 as shown in FIG. 9D.

For the sub-categories, probabilities obtained from the content pertaining to the same sub-category are summed to form the probability for the new program being in that sub-category. At the sub-category level, the same method is applied to  
30 compute the probability of a program being from the given category. The three levels of the program classification system; the category, the sub-category and the content, are used by the characterization processing 800 to form the characteristics vectors.

The program characteristics vectors 150 in general are represented in FIGS. 9A through 9F. FIGS. 9A, 9B and 9C are an example of deterministic vectors. This set of vectors is generated when the program characteristics are well defined, as  
5 can occur when the source related text 136 or the EPG 140 contains specific fields identifying the category and the sub-category. A program rating can also provided by the EPG 140.

In the case that these characteristics are not specified, a statistical set of vectors is generated from the process  
10 described in accordance with FIG. 8. FIGS. 9D-9F illustrate the probability that a program being watched is from the given category and program classification, respectively.

FIG. 10A illustrates sets of logical heuristic rules which form part of the heuristic rules 160. In a preferred  
15 embodiment, logical heuristic rules are obtained from sociological or psychological studies. Two types of rules are illustrated in FIG. 10A. The first type links an individual's viewing characteristics to demographic characteristics such as gender, age, and income level. A channel changing rate rule  
20 1030 attempts to determine gender based on channel change rate. An income related channel change rate rule 1010 attempts to link channel change rates to income brackets. A second type of rules links particular programs to particular audience, as illustrated by a gender determining rule 1050 which links the  
25 program category/sub-category with a gender. The result of the application of the logical heuristic rules illustrated in FIG. 10A are probabilistic determinations of factors including gender, age, and income level. Although a specific set of logical heuristic rules has been used as an example, a wide  
30 number of types of logical heuristic rules can be used to realize the present invention. In addition, these rules can be changed based on learning within the system or based on external studies which provide more accurate rules.

FIG. 10B illustrates a set of the heuristic rules

expressed in terms of conditional probabilities. In the example shown in FIG. 10B, each of several categories have associated conditional probabilities for demographic factors such as age, income, family size and gender composition.

5        FIG. 11 illustrates an entity-relationship diagram for the generation of the demographic vectors 170. In a preferred embodiment, the heuristic rules 160 are applied along with the characteristic vectors 150 in a target analysis process 1100 to form the demographic vectors 170. The characteristic vectors  
10    150 indicate a particular aspect of a program, such as its violence level. The heuristic rules 160 indicate that a particular demographic group has a preference for that program. As an example, it may be the case that young males have a higher preference for violent programs than other sectors of  
15    the population. Thus, a program which has the characteristic vectors 150 indicating a high probability of having violent content, when combined with the heuristic rules 160 indicating that "young males like violent programs," will result, through the target analysis process 1100, in the demographic vectors  
20    170 which indicate that there is a high probability that the program is being watched by a young male.

      The target analysis process 1100 can be realized using software programmed in a variety of languages which processes mathematically the heuristic rules 160 to derive the  
25    demographic vectors 170. The table representation of the heuristic rules 160 illustrated in FIG. 10B expresses the probability that the individual or household is from a specific demographic group based on a program with a particular category. This can be expressed, using probability terms as  
30    follow "the probability that the individuals are in a given demographic group conditional to the program being in a given category".

      Expressing the probability that a program is destined to a specific demographic group can be determined by applying Bayes

rule. This probability is the sum of the conditional probabilities that the demographic group likes the program, conditional to the category 144 weighted by the probability that the program is from that category 144. In a preferred embodiment, the program target analysis can calculate the program demographic vectors by the application of logical heuristic rules 160, as illustrated in FIG. 10A, and by the application of heuristic rules expressed as conditional probabilities as shown in FIG. 10B. Logical heuristic rules can be applied using logical programming and fuzzy logic using techniques well understood by those skilled in the art, and are discussed in the text by S. V. Kartalopoulos entitled "Understanding Neural Networks and Fuzzy Logic."

Conditional probabilities can be applied by simple mathematical operations multiplying program context vectors by matrices of conditional probabilities. By performing this process over all the demographic groups, the target analysis process 1100 can measure how likely a program is to be of interest to each demographic group. Those probabilities values form the demographic vector 170 represented in FIG. 12.

As an example, the heuristic rules expressed as conditional probabilities shown in FIG. 10B are used as part of a matrix multiplication in which the program characteristics vector 150 of dimension N, such as those shown in FIGS. 9A-9F is multiplied by an N x M matrix of heuristic rules expressed as conditional probabilities, such as that shown in FIG. 10B. The resulting vector of dimension M is a weighted average of the conditional probabilities for each category and represents the household demographic characteristics 190. Similar processing can be performed at the sub-category and content levels.

FIG. 13 illustrates an entity-relationship diagram for the generation of household session demographic data 1310 and a household session interest profile 1320. In a preferred



embodiment, the actual subscriber selection data 199 is used along with the program characteristics vectors 150 in a session characterization process 1300 to generate the household session interest profile 1320. The actual subscriber selection data 5 199 is based upon one or more advertisement selection records and indicates which advertisements the subscriber is watching, for how long and at what volume they are watching. Similarly, the actual subscriber selection data 199 also indicates which programs subscriber is watching, for how long and at what 10 volume they are watching the program.

In a preferred embodiment, the session characterization process 1300 forms a weighted average of the program characteristics vectors 150 in which the time duration the program is watched is normalized to the session time (typically 15 defined as the time from which the unit was turned on to the present). The program characteristics vectors 150 are multiplied by the normalized time duration (which is less than one unless only one program has been viewed) and summed with the previous value. Time duration data, along with other 20 subscriber viewing information, is available from the subscriber selection data 110A and 110B. The resulting weighted average of program characteristics vectors 150 forms the household session interest profile 1320, with each program contributing to the household session interest profile 1320 25 according to how long it was watched. The household session interest profile 1320 is normalized to produce probabilistic values of the household programming interests during that session.

In an alternate embodiment, the heuristic rules 160 are 30 applied to both the actual subscriber selection data 199 and the characteristics vectors 150 to generate the household session demographic data 1310 and the household session interest profile 1320. In this embodiment, weighted averages of the characteristics vectors 150 are formed based on the

actual subscriber selection data 199, and the heuristic rules 160 are applied. In the case of logical heuristic rules as shown in FIG. 10A, logical programming can be applied to make determinations regarding the household session demographic data 1310 and the household session interest profile 1320. In the case of heuristic rules in the form of conditional probabilities such as those illustrated in FIG. 10B, a dot product of the time averaged values of the program characteristics vectors 150 can be taken with the appropriate matrix of heuristic rules 160 to generate both the household session demographic data 1310 and the household session interest profile 1320.

Volume control measurements which form part of the subscriber selection data 110A and 110B can also be applied in the session characterization process 1300 to form a household session interest profile 1320. This can be accomplished by using normalized volume measurements in a weighted average manner similar to how time duration is used. Thus, muting a show results in a zero value for volume, and the program characteristics vector 150 for this show will not be averaged into the household session interest profile 1320.

FIG. 14 illustrates an entity-relationship diagram for the generation of average household demographic characteristics 190 and session household demographic characteristics 190. A household demographic characterization process 1400 generates the household demographic characteristics 190 represented in table format in FIG. 15. The household demographic characterization process 1400 uses the household viewing habits 195 in combination with the heuristic rules 160 to determine demographic data. For example, a household with a number of minutes watched of zero during the day may indicate a household with two working adults. Both logical heuristic rules as well as rules based on conditional probabilities can be applied to the household viewing habits 195 to obtain the household

demographics characteristics 190.

The household viewing habits 195 is also used by the system to detect out-of-habits events. For example, if a household with a zero value for the minutes watched column 702 at late night presents a session value at that time via the household session demographic data 1310, this session will be characterized as an out-of-habits event and the system can exclude such data from the average if it is highly probable that the demographics for that session are greatly different than the average demographics for the household. Nevertheless, the results of the application of the household demographic characterization process 1400 to the household session demographic data 1310 can result in valuable session demographic data, even if such data is not added to the average demographic characterization of the household.

FIG. 15 illustrates the average and session household demographic characteristics 190. A household demographic parameters column 1501 is followed by an average value column 1505, a session value column 1503, and an update column 1507. The average value column 1505 and the session value column 1503 are derived from the household demographic characterization process 1400. The deterministic parameters such as address and telephone numbers can be obtained from an outside source or can be loaded into the system by the subscriber or a network operator at the time of installation. Updating of deterministic values is prevented by indicating that these values should not be updated in the update column 1507.

FIG. 16 illustrates an entity-relationship diagram for the generation of the household interest profile 180 in a household interest profile generation process 1600. In a preferred embodiment, the household interest profile generation process 1600 comprises averaging the household session interest profile 1320 over multiple sessions and applying the household viewing habits 195 in combination with the heuristic rules 160 to form

the household interest profile 180 which takes into account both the viewing preferences of the household as well as assumptions about households/subscribers with those viewing habits and program preferences.

5        FIG. 17 illustrates an exemplary household interest profile 180 that includes a programming types row 1709, a products types row 1707, a household interests column 1701, an average value column 1703, and a session value column 1705.

10        The product types row 1707 gives an indication as to what type of advertisement the household would be interested in watching, thus indicating what types of products could potentially be advertised with a high probability of the advertisement being watched in its entirety. The programming types row 1709 suggests what kind of programming the household  
15        is likely to be interested in watching. The household interests column 1701 specifies the types of programming and products which are statistically characterized for that household.

20        As an example of the industrial applicability of the invention, a household will perform its normal viewing routine without being requested to answer specific questions regarding likes and dislikes. Children may watch television in the morning in the household, and may change channels during commercials, or not at all. The television may remain off  
25        during the working day, while the children are at school and day care, and be turned on again in the evening, at which time the parents may "surf" channels, mute the television during commercials, and ultimately watch one or two hours of broadcast programming. The present invention provides the ability to  
30        characterize the household, and may make the determination that there are children and adults in the household, with program and product interests indicated in the household interest profile 180 corresponding to a family of that composition. A household with two retired adults will have a completely

different characterization which will be indicated in the household interest profile 180.

Since a viewing session is likely to be dominated by a particular viewer, the session values may, in some  
5 circumstances, correspond most closely to the subscriber values, while the average values may, in some circumstances, correspond most closely to the household values. The average value being a time average of data, where the averaging period may be several days, weeks, months, or the time between resets  
10 of unit.

Although the present invention has been largely described in the context of a single computing platform receiving programming, the AMFS 100 can be realized as part of a client-server architecture, as illustrated in FIG. 18. Residence 1800  
15 contains a personal computer (PC) 1820 as well as the combination of a television 1810 and a set-top box 1808, which can request and receive programming. The equipment in residence 1800, or similar equipment in a small or large business environment, forms the client side of the network as  
20 defined herein. Programming is delivered over an access network 1830, which may be a cable television network, telephone type network, or other access network. Information requests are made by the client side to a server 1840 which forms the server side of the network. The server 1840 has  
25 content locally which it provides to the subscriber, or requests content on behalf of the subscriber from a third party content provider 1860, as illustrated in FIG. 18. Requests made on behalf of the client side by server 1840 are made across a wide area network 1850 which can be the Internet or  
30 other public or private network. Techniques for making requests on behalf of a client are frequently referred to a proxy techniques and are well known to those skilled in the art. The server side receives the requested programming which

is displayed on PC 1820 or television 1810 according to which device made the request.

According to one embodiment, the server side maintains the advertising selection data 110A and subscriber selection data 110B which it is able to compile based on its operation as a proxy for the client side. Retrieval of source related information 130, the program target analysis process 1100, the program characterization process 800, the session characterization process 1300, the household demographic characterization process 1400, and the household interest profile generation process 1600 can be performed by server 1840.

Referring to FIG. 19 an advertisement monitoring table is illustrated, in which an advertisement ID (AD ID) column 1915 contains a numerical ID for an advertisement which was transmitted with the advertisement in the form of a Program ID, http address, or other identifier which is uniquely associated with the advertisement. A product column 1921 contains a product description which indicates the type of product that was advertised. A brand column 1927 indicates the brand name of the product or can alternatively list a generic name for that product. A percent watched column 1933 indicates the percentage of the advertisement the subscriber viewed. In an alternate embodiment, a letter rating or other type of rating is used to indicate the probability that the advertisement was watched. A volume column 1937 indicates the volume level at which the advertisement was watched.

As an example of the industrial applicability of the invention, a manufacturer may develop an advertising strategy which includes the insertion of advertisements during popular evening programs. The costs for such ad insertions can be extremely high. In order to insure the cost effectiveness of this advertising strategy, the manufacturer has the advertisements placed during less watched but similar programs

and monitors how subscribers react, and can determine approximately how many times the advertisement has been watched out of all of the possible viewings. This data can be used to confirm the potential effectiveness of the advertisement and to subsequently determine if purchasing the more expensive time during evening programming will be cost-effective, or if the advertisement should be modified or placed in other programming.

Continuing this example, the manufacturer may place an advertisement for viewing during "prime time" for an initial period but can subsequently cancel broadcasts of the advertisement if it is found that the majority of subscribers never see the advertisement.

Although this invention has been illustrated by reference to specific embodiments, it will be apparent to those skilled in the art that various changes and modifications may be made which clearly fall within the scope of the invention. The invention is intended to be protected broadly within the spirit and scope of the appended claims.

20

**Claims**

What is claimed is:

1. A method for generating a subscriber profile based on advertisements watched by the subscriber, the method  
5 comprising:  
  
monitoring subscriber advertising viewing activities;  
  
collecting advertisement selection data from the monitored subscriber advertising viewing activities; and  
  
generating a subscriber profile based on the advertisement  
10 selection data.  
  
2. The method of claim 1, further comprising retrieving advertisement related information for advertisements identified in the advertisement selection data wherein the advertisement  
15 related information contains descriptive fields related to the advertisements.  
  
3. The method of claim 2, wherein said retrieving includes context mining text associated with the  
20 advertisements.  
  
4. The method of claim 3, wherein the text is derived from closed-captioning data associated with the advertisement.



5. The method of claim 3, wherein the text includes a product name field.

6. The method of claim 3, wherein the text includes a product brand field.

7. The method of claim 1, wherein said monitoring subscriber advertising viewing activities includes monitoring volume levels.

10

8. The method of claim 1, wherein said monitoring subscriber advertising viewing activities includes monitoring channel change requests by the subscriber.

15 9. The method of claim 1, wherein said collecting advertisement selection data includes determining the extent to which an advertisement is viewed by the subscriber.

10. The method of claim 1, wherein the subscriber profile is a subscriber product interest profile.

11. The method of claim 1, further includes:  
monitoring subscriber program viewing activities;

collecting program selection data for the monitored subscriber program viewing activities;

combining the program selection data with the advertisement selection data;

5       evaluating the combined program selection data and the advertisement selection data to filter out irrelevant data and generate a record of actual subscriber selection data; and

processing the actual subscriber selection data to create the subscriber profile.

10

12. The method of claim 11, wherein said monitoring subscriber program viewing activities includes monitoring viewing time durations for selected source material.

15       13. The method of claim 11, wherein said evaluating the combined program selection data includes evaluating channel change commands and associated viewing times.

14. The method of claim 13, wherein said evaluating the  
20 combined program selection data includes filtering out any channel change commands if the associated viewing times are below a pre-determined threshold.

15. The method of claim 14, wherein the filtered out channel change commands correspond to channel surfing activities.

5        16. The method of claim 14, wherein the filtered out channel change commands correspond to channel jumping activities.

17. The method of claim 11, wherein said evaluating  
10 includes evaluating viewing times and filtering out any viewing periods if no subscriber activity has been received within a pre-determined period of time.

18. The method of claim 17, wherein the filtered out  
15 viewing periods correspond to dead periods implying that the subscriber is not actively watching a television or other multimedia device.

19. The method of claim 11, wherein the subscriber  
20 profile is a program preference profile for the subscriber, the program preference profile indicating the type of programming of interest to the subscriber.

20. The method of claim 11, wherein the subscriber profile is a product preference profile for the subscriber, the product preference profile indicating the type of products of interest to the subscriber.

5

21. The method of claim 11, wherein the subscriber profile is an advertising preference profile for the subscriber, the advertising preference profile indicating the type of advertising of interest to the subscriber.

10

22. A data processing system for generating a subscriber profile based on advertisements watched by the subscriber, the system comprising:

a storage medium;

15 means for monitoring subscriber advertising activities;

means for collecting advertisement selection data for the monitored subscriber advertising activities;

means for retrieving advertisement related information for ads in the advertisement selection data, wherein the

20 advertisement related information contains descriptive fields corresponding to the advertisements;

means for determining the extent to which the advertisements are viewed by the subscriber; and

means for generating a subscriber profile based on the advertisement related information and the extent to which the advertisements are viewed.

5        23. The system of claim 22, wherein the means for monitoring subscriber activity includes means for monitoring volume levels wherein the volume levels correspond to subscriber selection volume levels.

10       24. The system of claim 22, wherein the means for monitoring subscriber activity includes means for monitoring channel change requests initiated by the subscriber.

15       25. The system of claim 22, wherein the means for retrieving advertisement related information includes means for context mining of textual information associated with selected source material.

20       26. The system of claim 25, wherein the textual information is text derived from closed-captioning data associated with the advertisement.

27. The system of claim 26, wherein the text derived from closed-captioning data associated with the advertisement includes a product name field.

5        28. The system of claim 26, wherein the text derived from closed-captioning data associated with the advertisement includes a product brand field.

29. The system of claim 22, further comprising:  
10        means for monitoring subscriber program viewing activities;  
         means for collecting program selection data for the monitored subscriber program viewing activities; and  
         means for combining the advertisement selection data and  
15        the program selection data.

30. The system of claim 29, further comprising:  
         means for filtering out irrelevant data and generating a record of actual subscriber selection data; and  
20        means for processing the actual subscriber selection data to create a subscriber profile.

31. The system of claim 22, wherein the subscriber profile is a program preference profile for the subscriber, the program preference profile indicating the type of programming of interest to the subscriber.

5

32. The system of claim 22, wherein the subscriber profile is a product preference profile for the subscriber, the product preference profile indicating the type of products of interest to the subscriber.

10

33. The system of claim 22, wherein the subscriber profile is an advertising preference profile for the subscriber, the advertising preference profile indicating the type of advertising of interest to the subscriber.

15

34. A client-server based data processing system for generating a subscriber profile based on the advertisements watched by a subscriber, the client-server based data processing system comprising:

20 computer processor means at a client side for receiving and displaying advertisements and transmitting channel change requests;

computer processor means at a server side for receiving the channel change requests and for processing data;

a storage medium;

means at the server side for monitoring subscriber activity including receiving subscriber channel change requests and storing subscriber channel change requests;

5 means at the server side for retrieving advertisement related information wherein the advertisement related information contains descriptive fields corresponding to an advertisement;

means at the server side for determining the extent to.  
10 which an advertisement is viewed by the subscriber;

means at the server side for storing the descriptive fields and the determination of the extent to which the advertisement is viewed by the subscriber; and

means at the server side for processing the descriptive  
15 fields and the extent to which the advertisement is viewed by the subscriber information to generate subscriber profile.

35. The system of claim 34, wherein the means for retrieving advertisement related information further comprises  
20 means for context mining of textual information associated with the selected source material.



36. The system of claim 35, wherein the textual information is text derived from closed-captioning data associated with the advertisement.

5        37. The system of claim 36, wherein the text derived from closed-captioning data associated with the advertisement includes a product name field.

10       38. The system of claim 36, wherein the text derived from closed-captioning data associated with the advertisement includes a product brand field.

15       39. The system of claim 34, further comprising:  
means for monitoring subscriber viewing activities related to actual programming;

means for collecting program selection data based on the actual programming; and

means for combining the advertisement selection data and the program selection data.

20

40. The system of claim 34, further comprising:

means for filtering out irrelevant data and generating a record of actual subscriber selection data; and

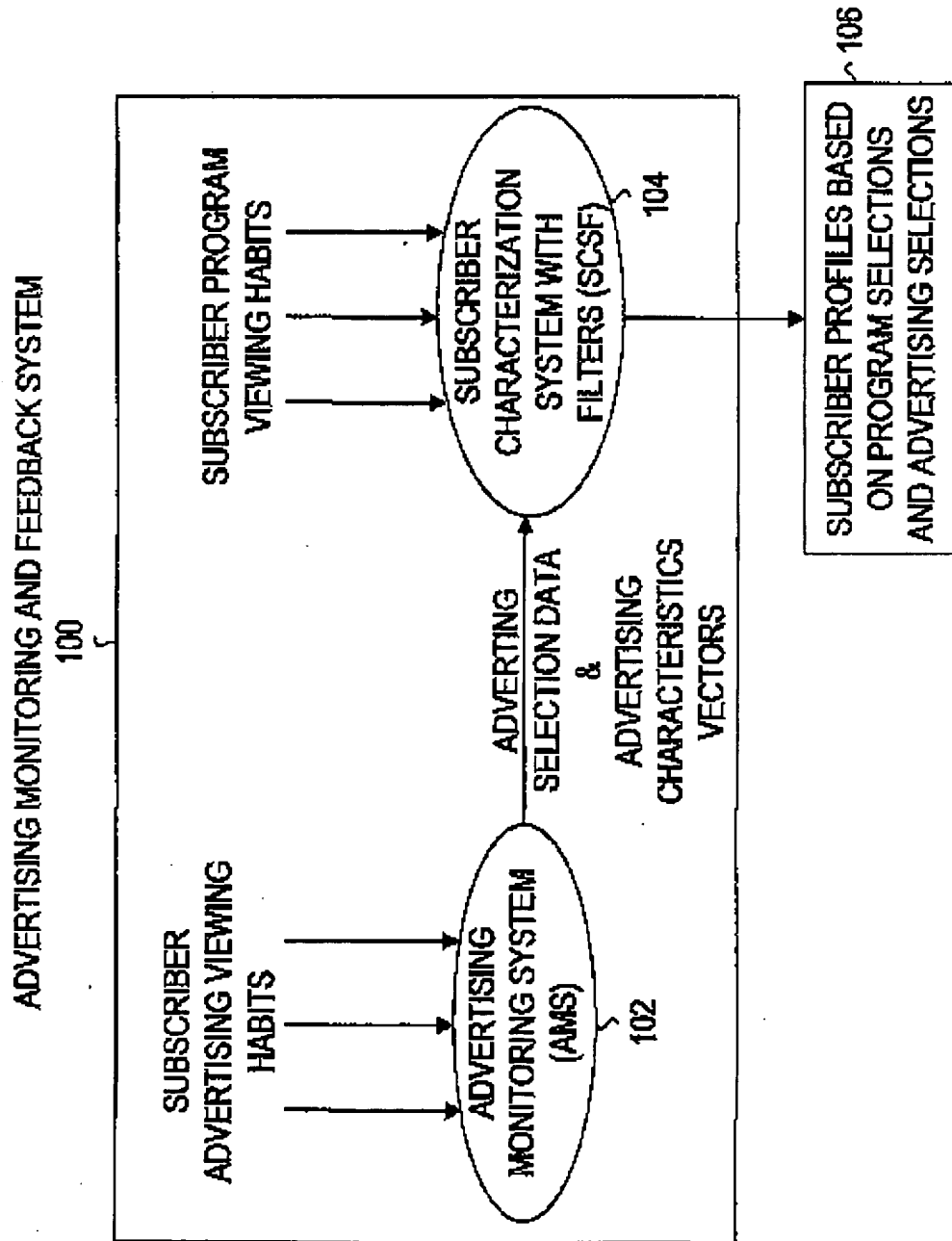
means for processing the actual subscriber selection data to create a subscriber profile.

41. The system described in claim 34, wherein the  
5 subscriber profile is a program preference profile for the subscriber, the program preference profile indicating the type of programming of interest to the subscriber.

42. The system of claim 34, wherein the subscriber  
10 profile is a product preference profile for the subscriber, the product preference profile indicating the type of products of interest to the subscriber.

43. The system of claim 34, wherein the subscriber  
15 profile is an advertising preference profile for the subscriber, the advertising preference profile indicating the type of advertising of interest to the subscriber.

1/27



2/27

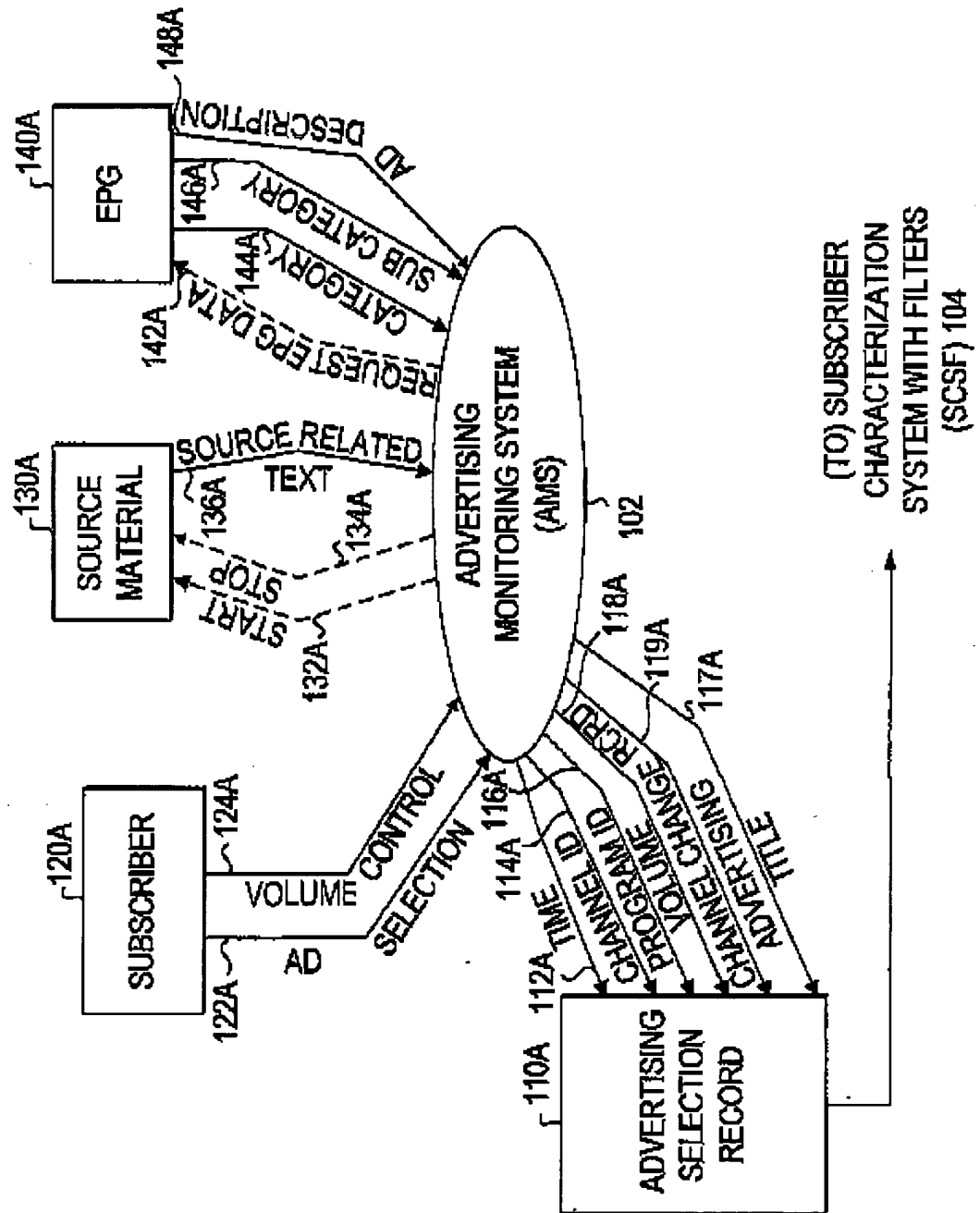
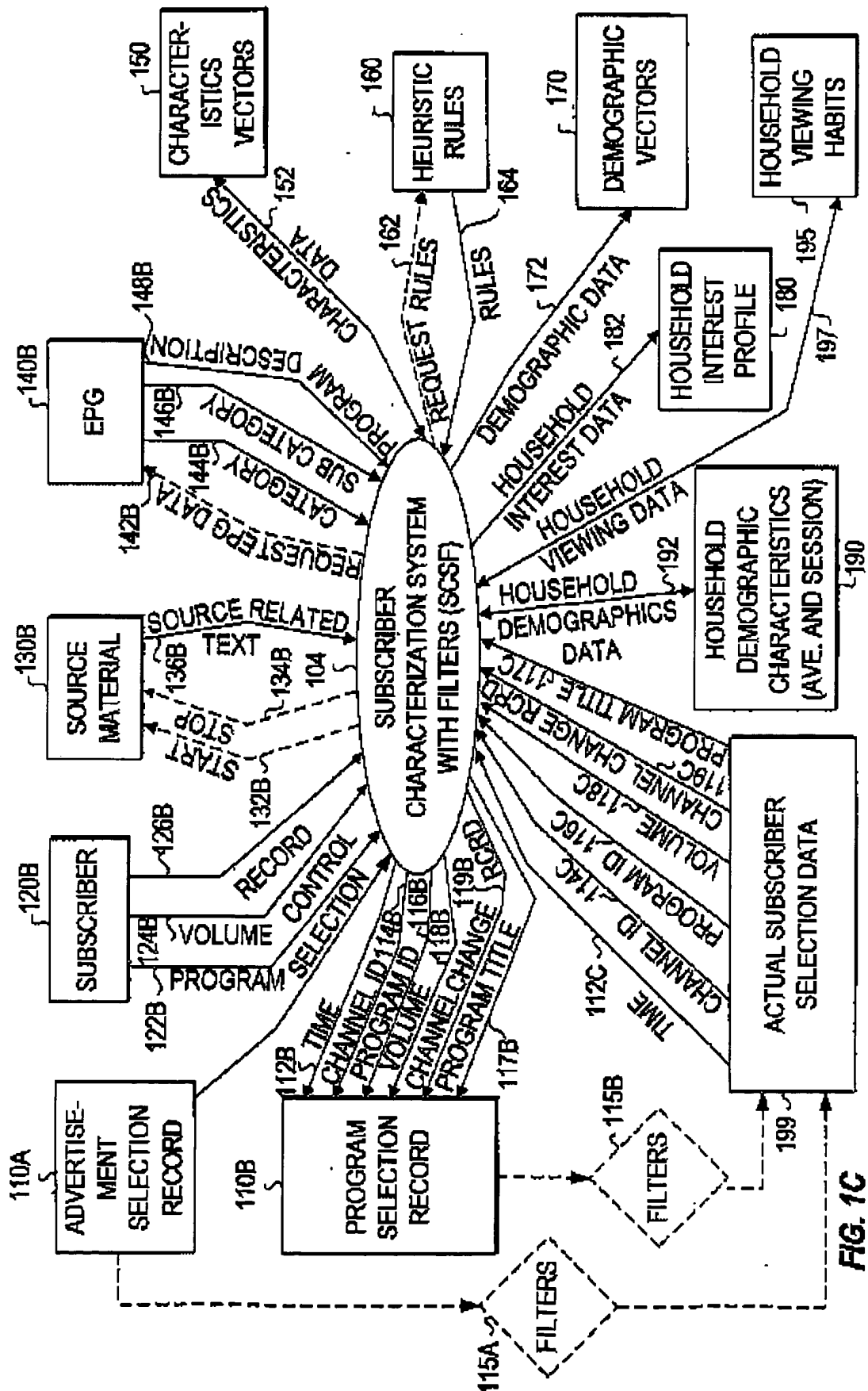
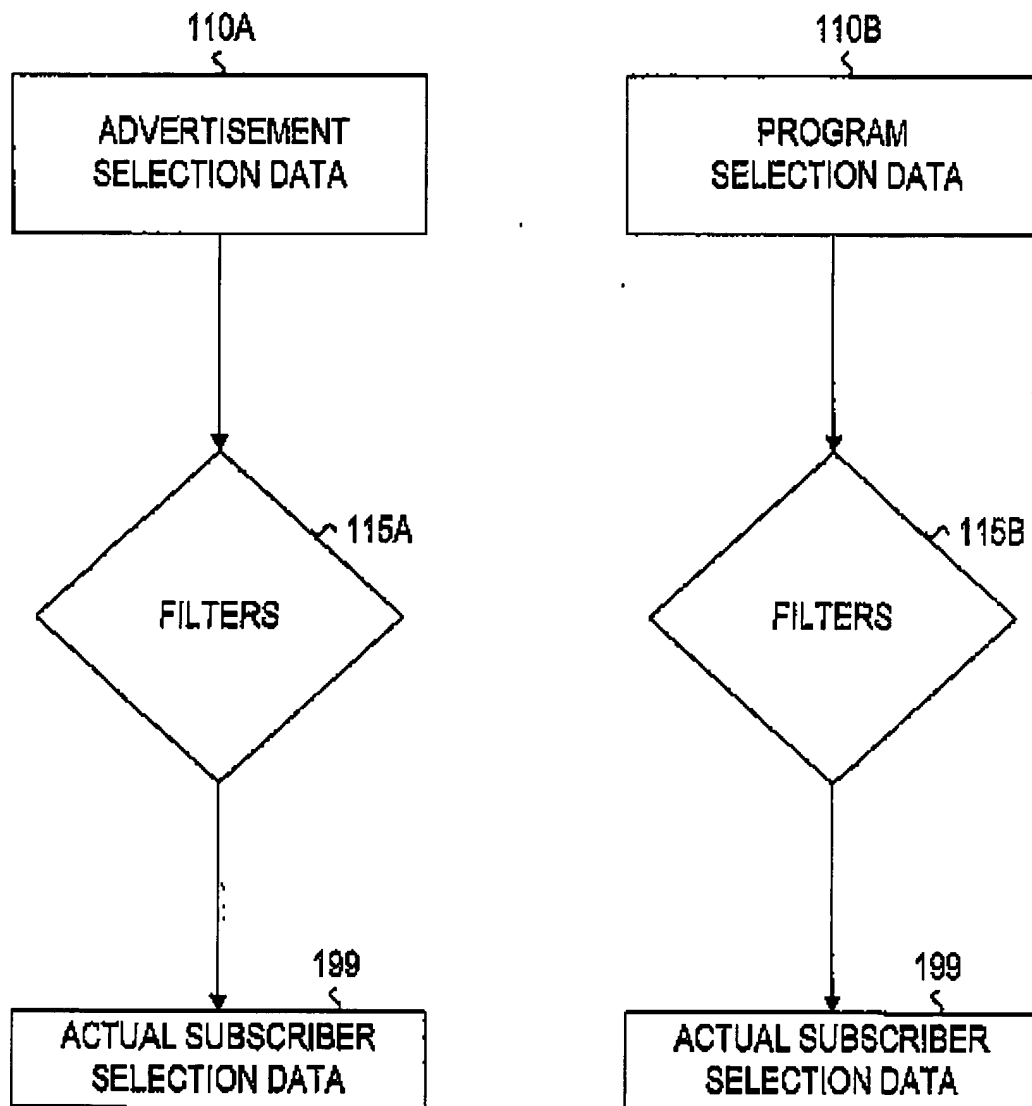


FIG. 1B

3/27



4/27

**FIG. 1D**

5/27

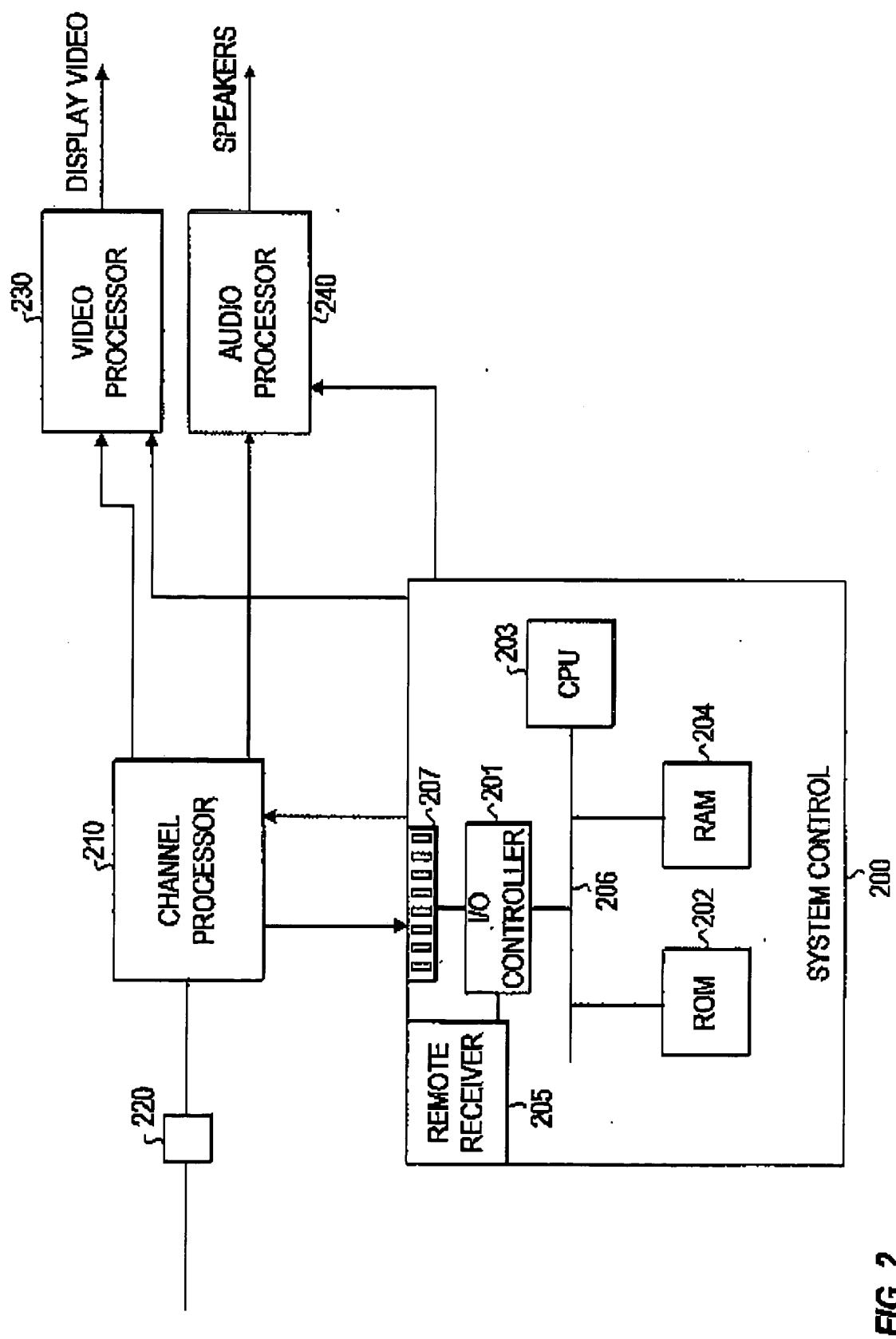


FIG. 2

6/27

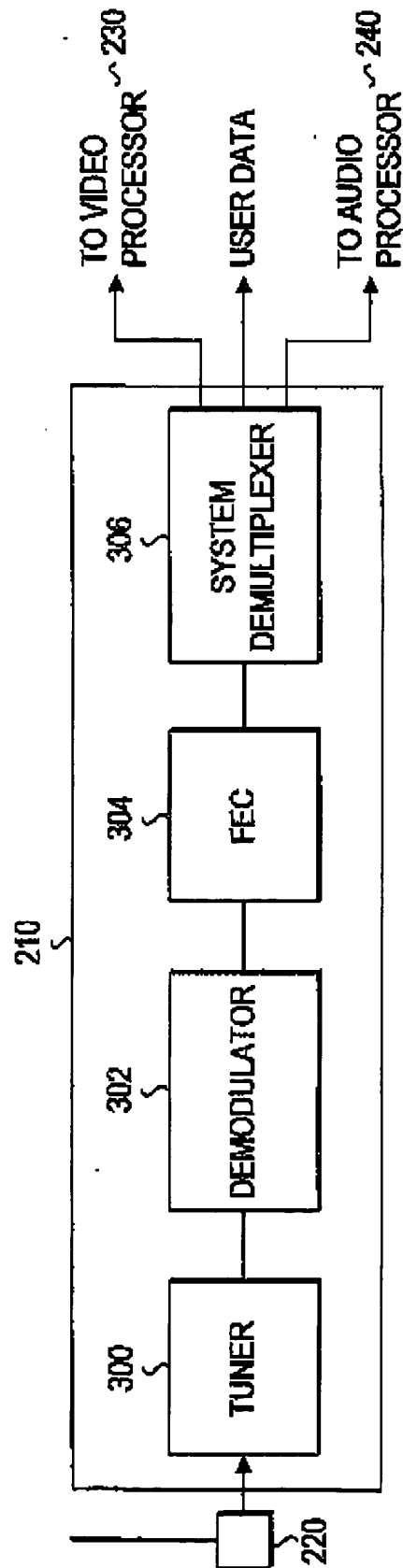


FIG. 3



7/27

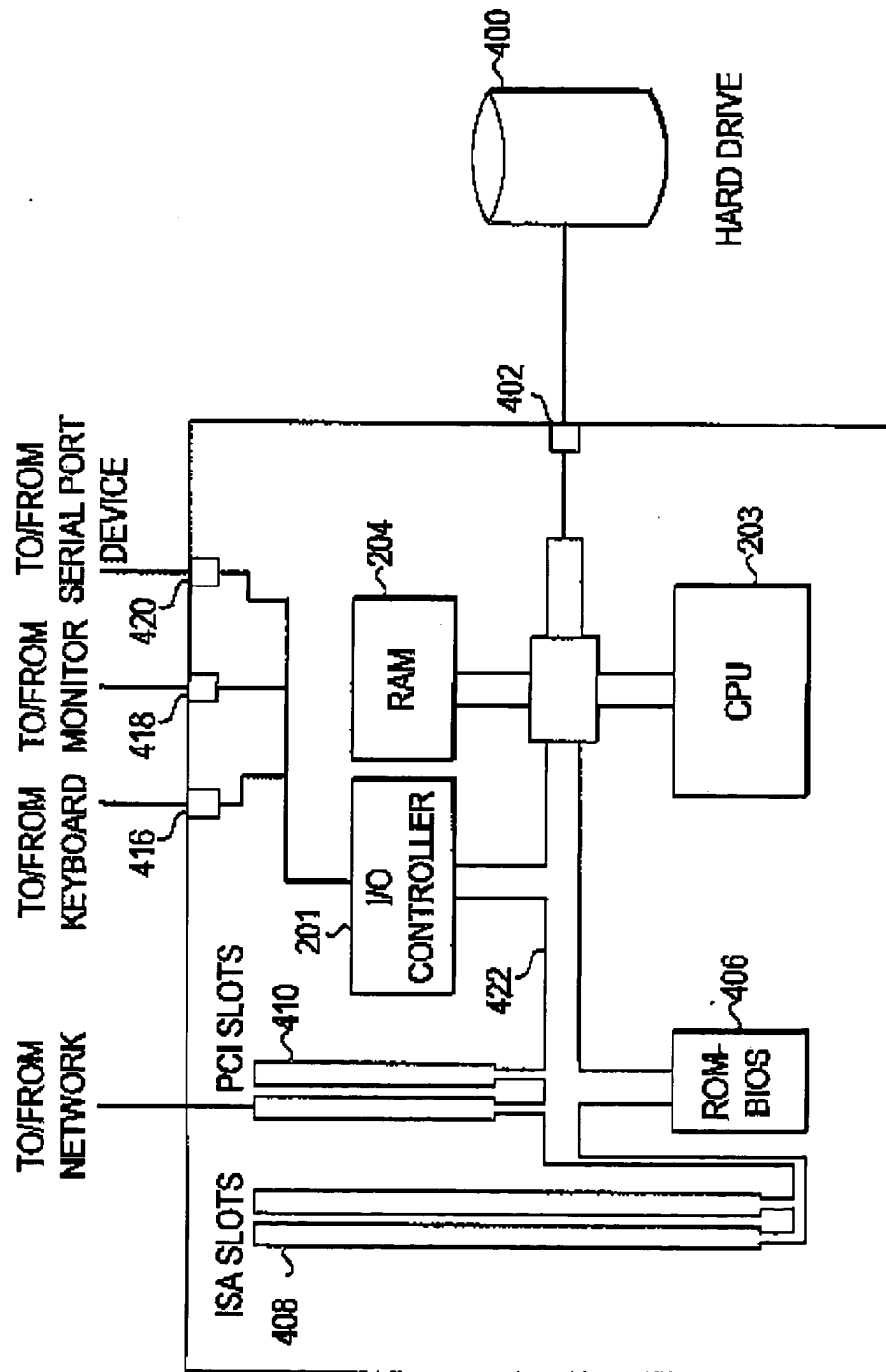


FIG. 4

8/27

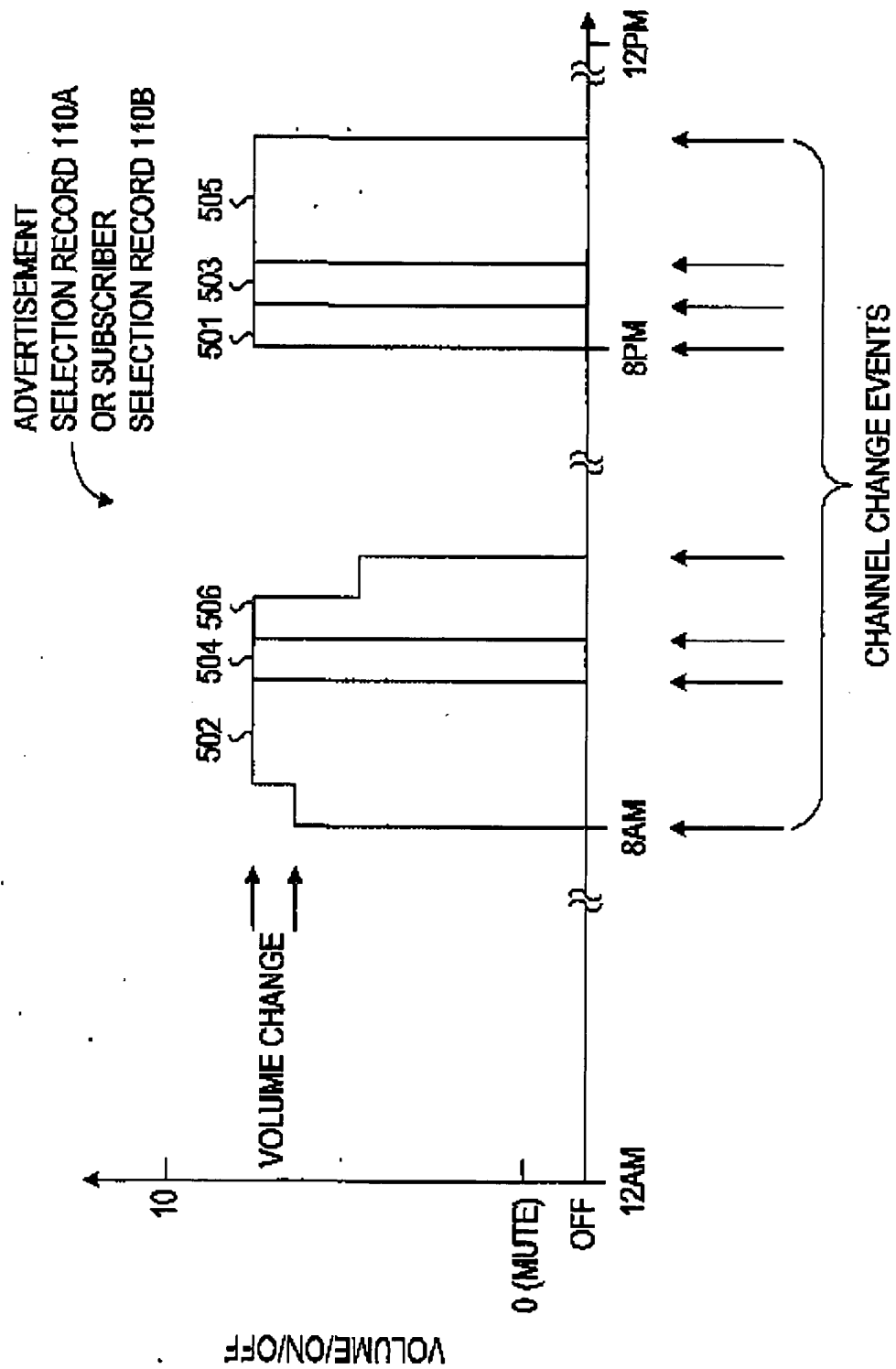


FIG. 5

9/27

602	604	603	601
TIME	CHANNEL ID	TITLE	VOLUME
08:01:25AM	06	"MORNING TV"	5/10
08:01:45AM	13		5/10
08:03:25AM	13		6/10
⋮			
06:11:25PM	09	"SEINFELD" "ADVERTISING" "SEINFELD" "ADVERTISING" "LIVING SINGLE"	5/10
06:15:23PM	09		5/10
06:17:25PM	09		5/10
06:28:10PM	09		5/10
06:30:07PM	52		5/10
⋮			

FIG. 6A

10/27

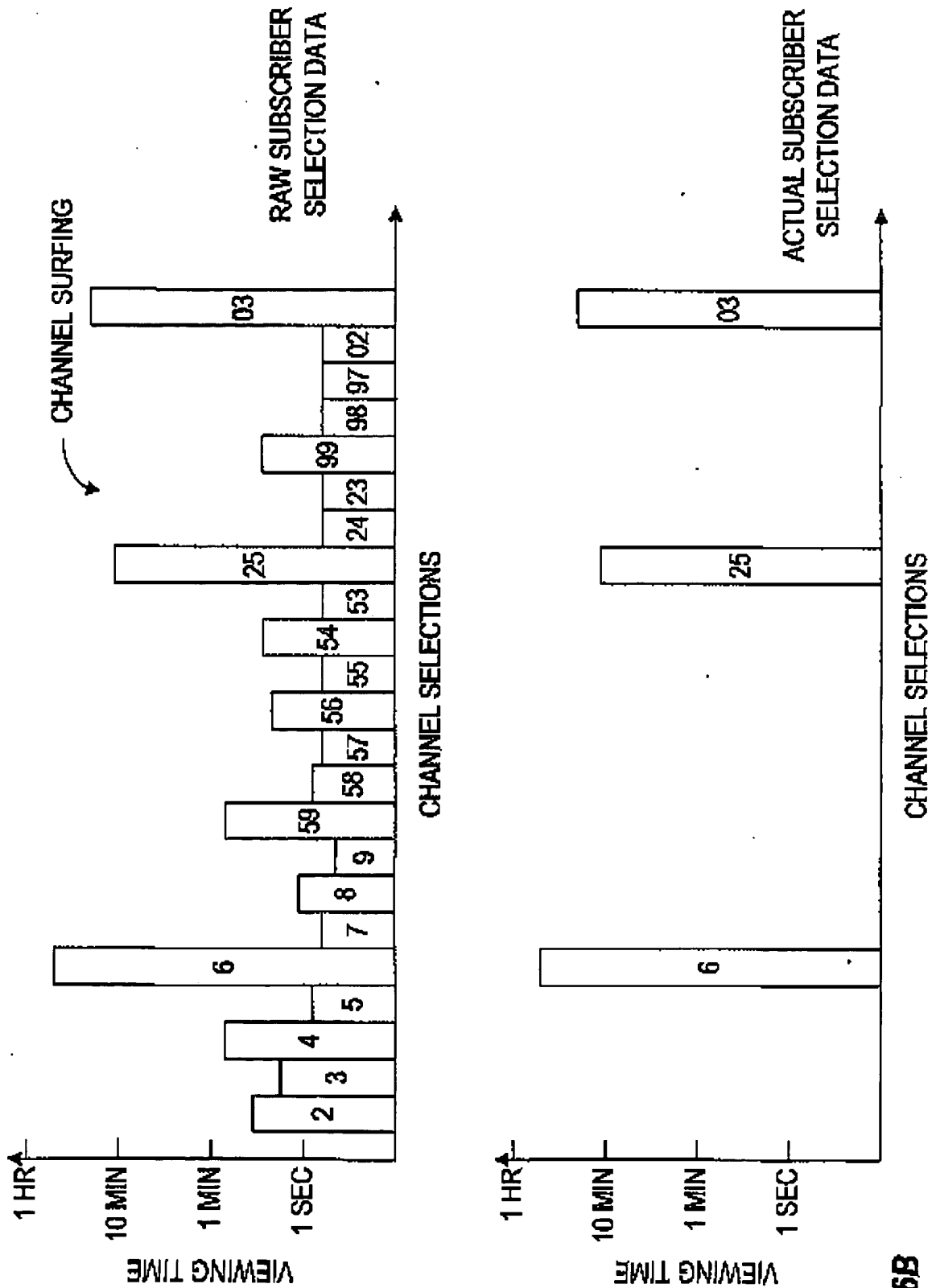


FIG. 6B

11/27

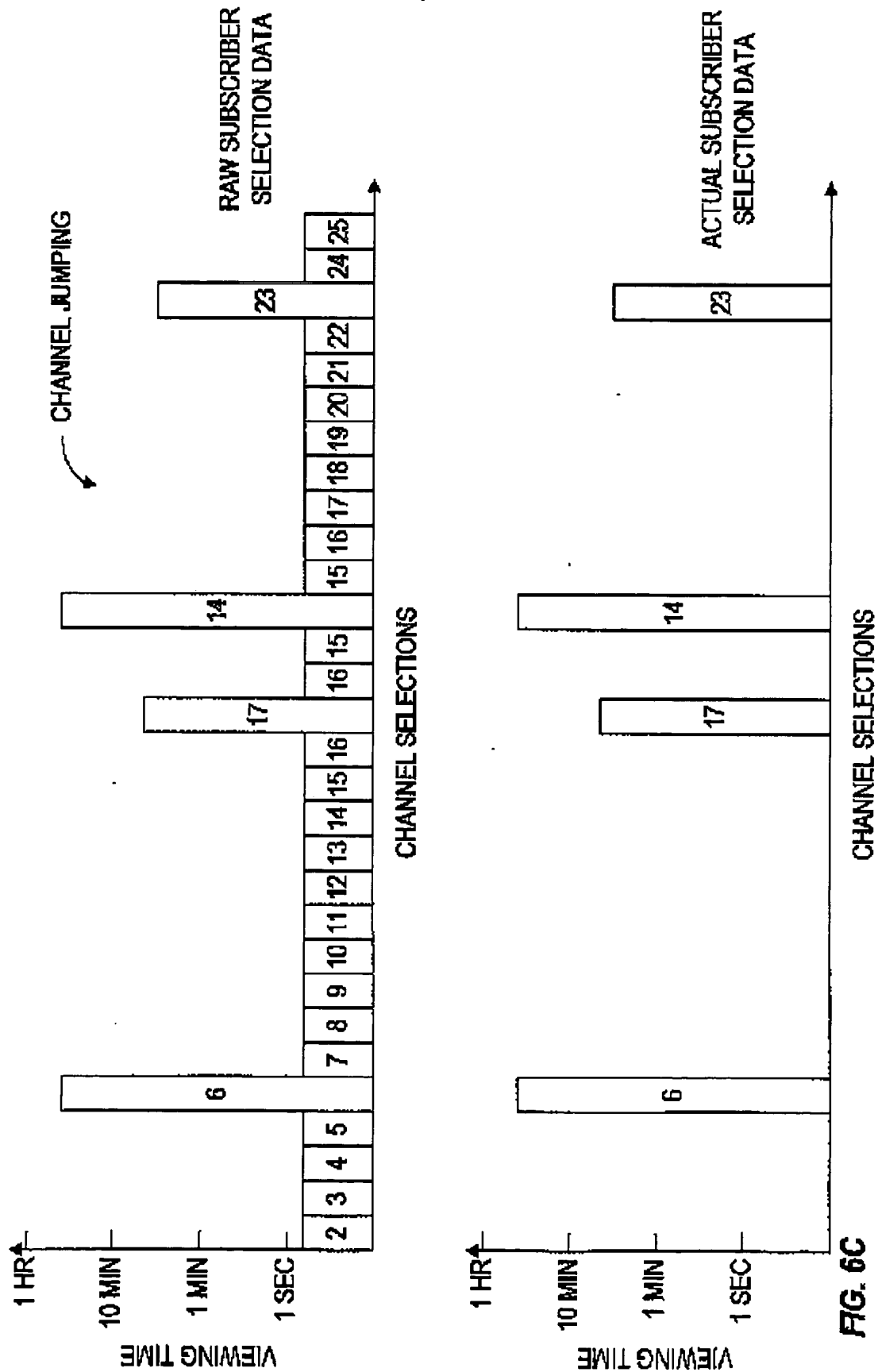


FIG. 6C

12/27

700 TIME OF DAY	702 MINUTES WATCHED	704 CHANNEL CHANGES	706 AVERAGE VOLUME
MORNING (6AM-9AM)	61	2	5/10
MID-DAY (9AM-3PM)	0	0	-
AFTERNOON (3PM-6PM)	0	0	-
NIGHT (6PM-10PM)	122	4	6/10
LATE NIGHT (12AM-6AM)	0	0	-
TOTAL	183	6	5.7/10

FIG. 7

13/27

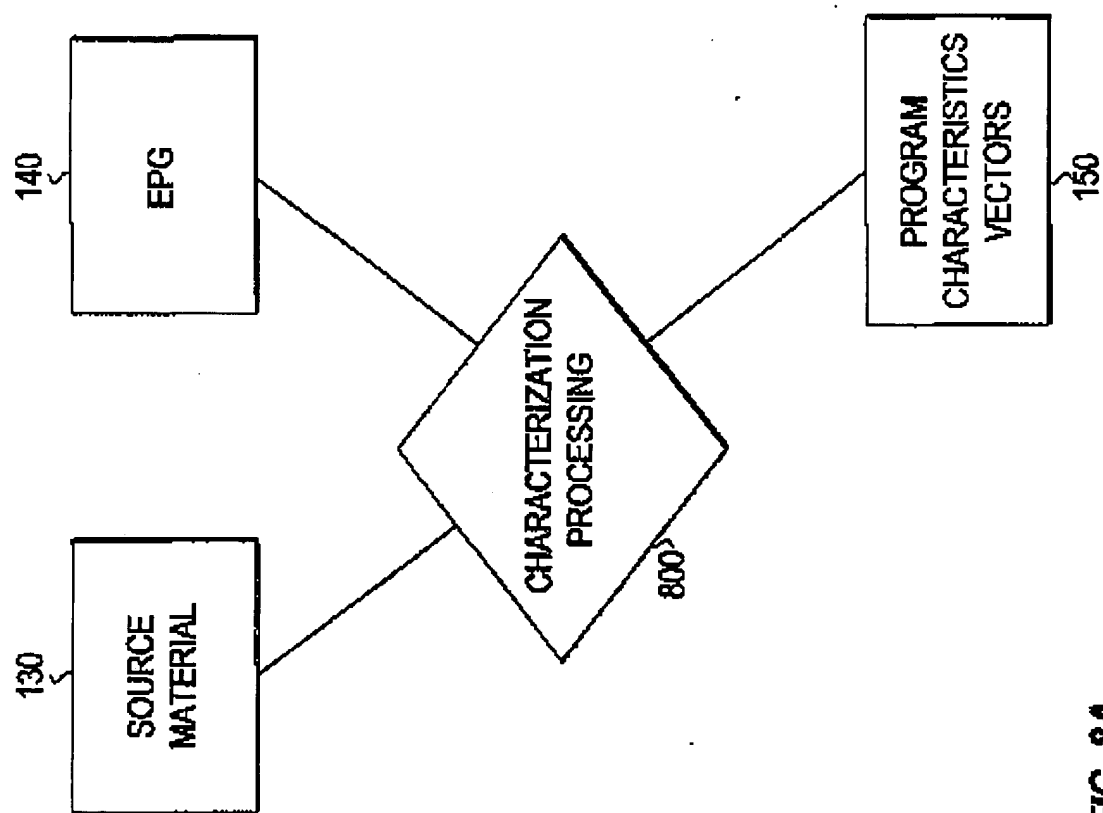


FIG. 8A

14/27

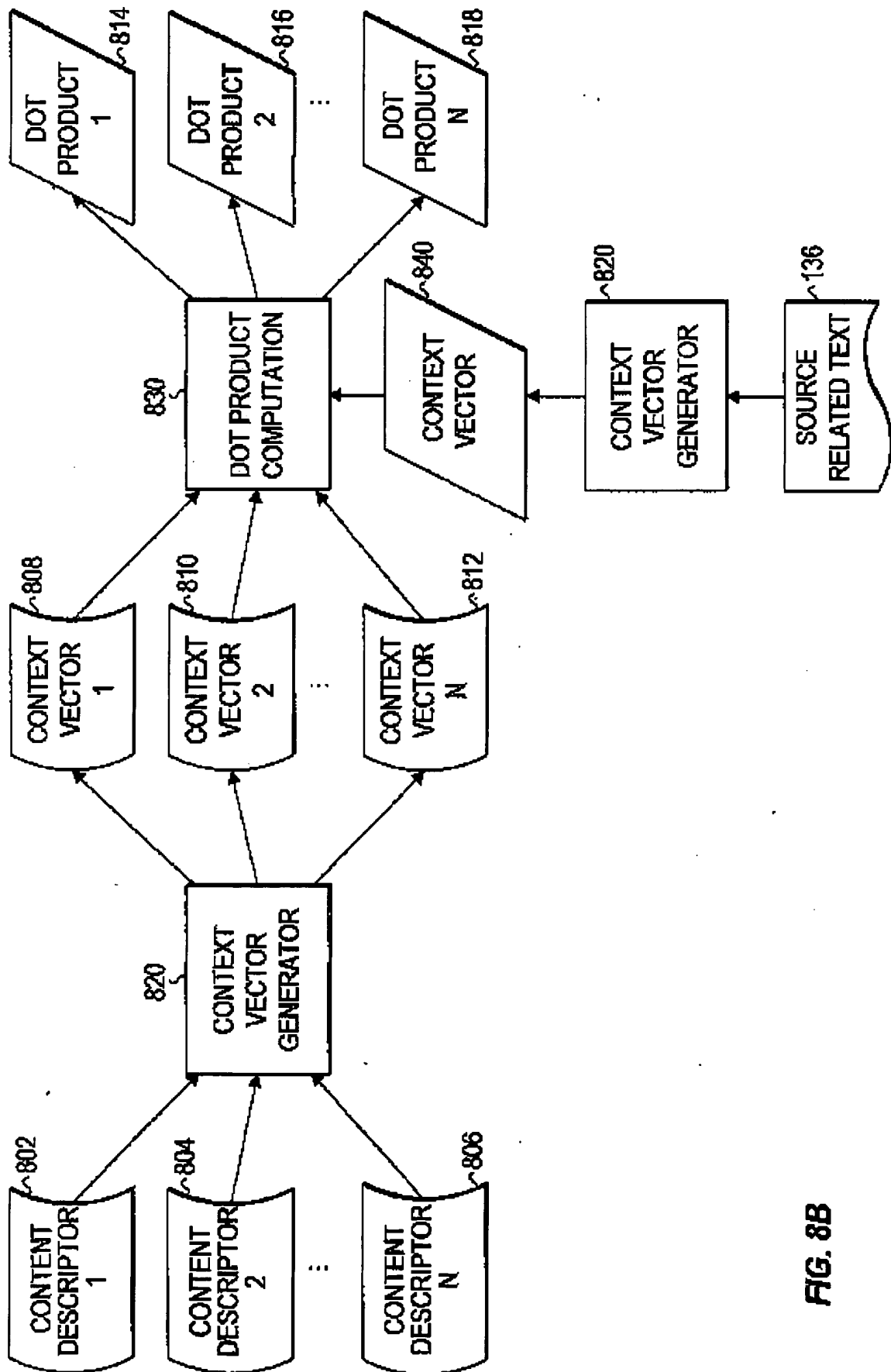


FIG. 8B



15/27

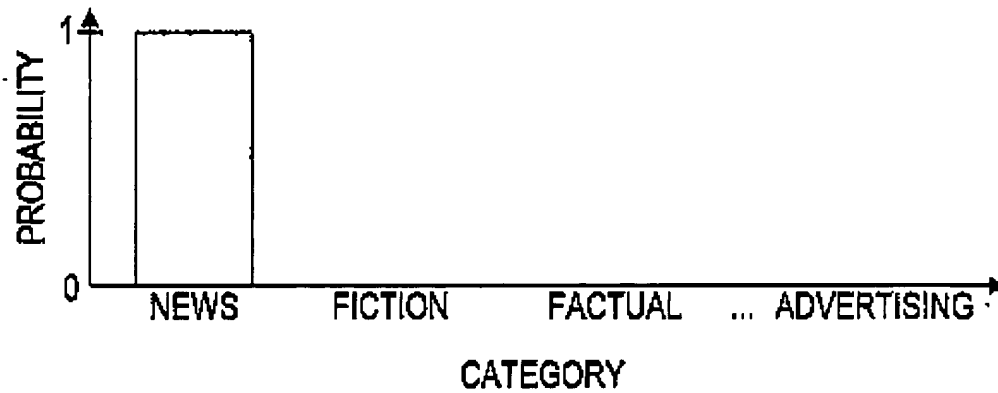


FIG. 9A

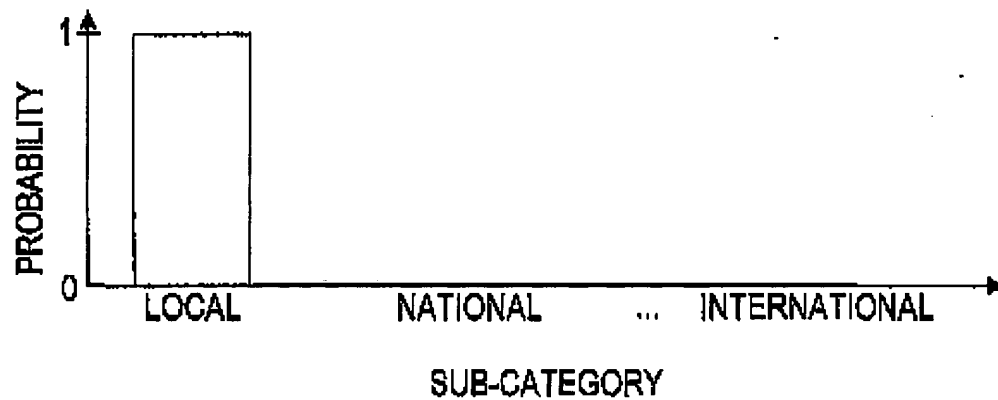


FIG. 9B

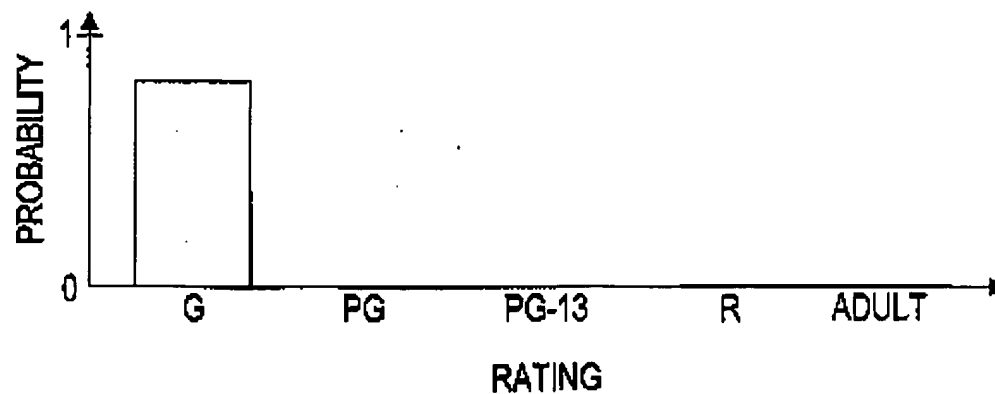
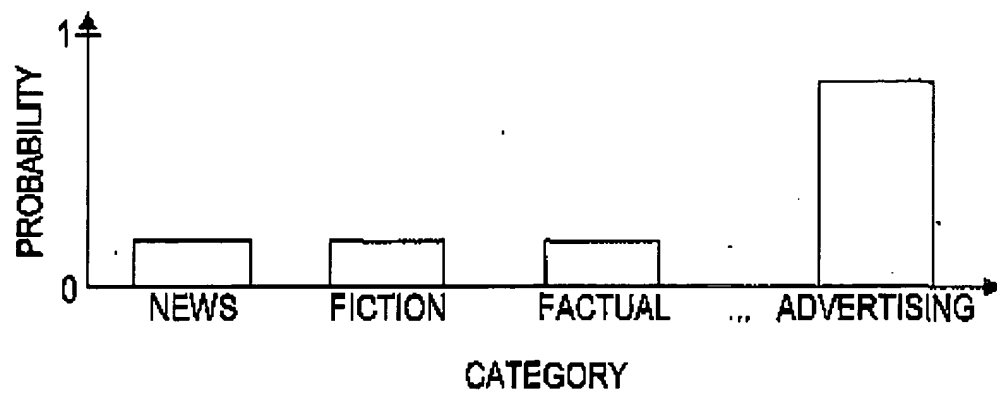
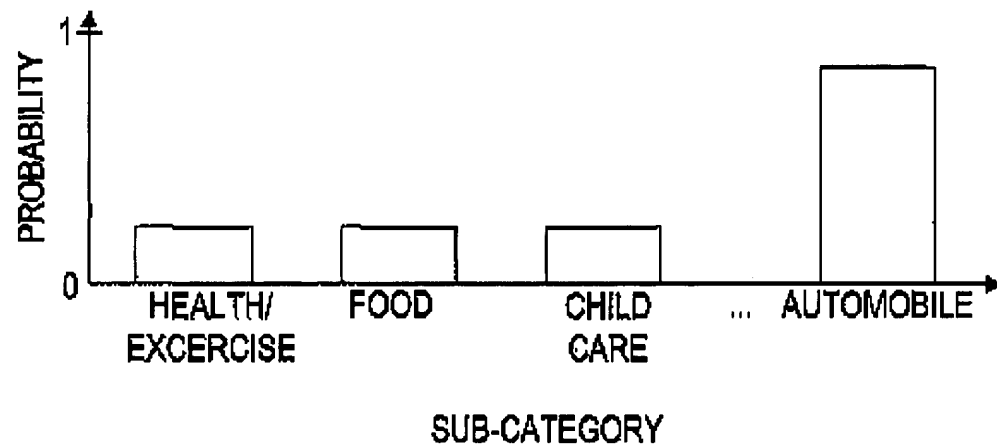
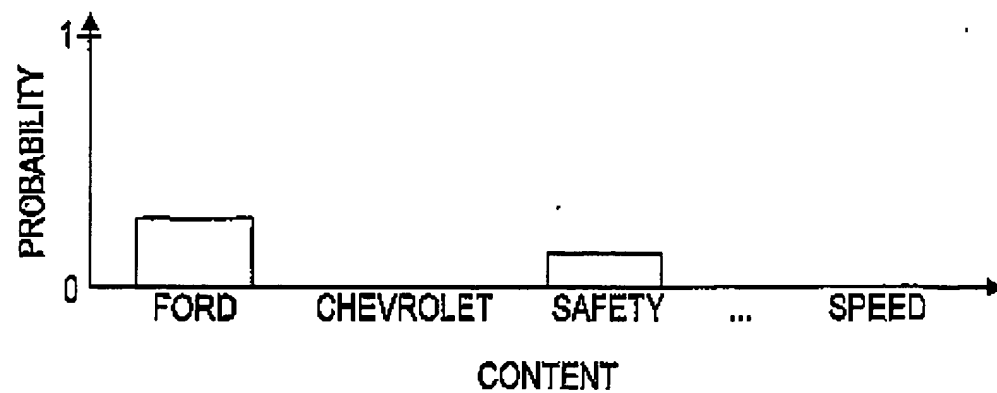


FIG. 9C

16/27

**FIG. 9D****FIG. 9E****FIG. 9F**

17/27

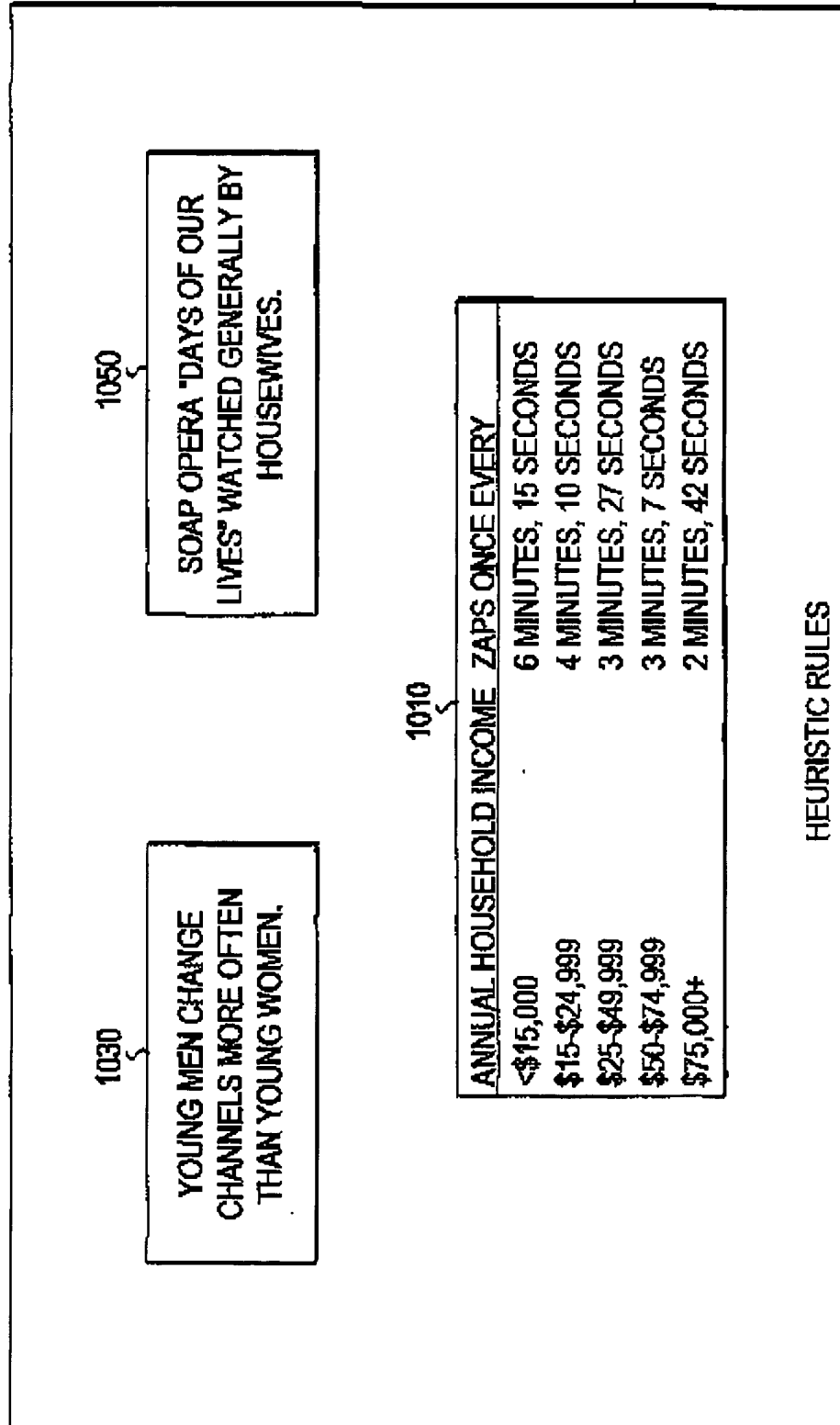


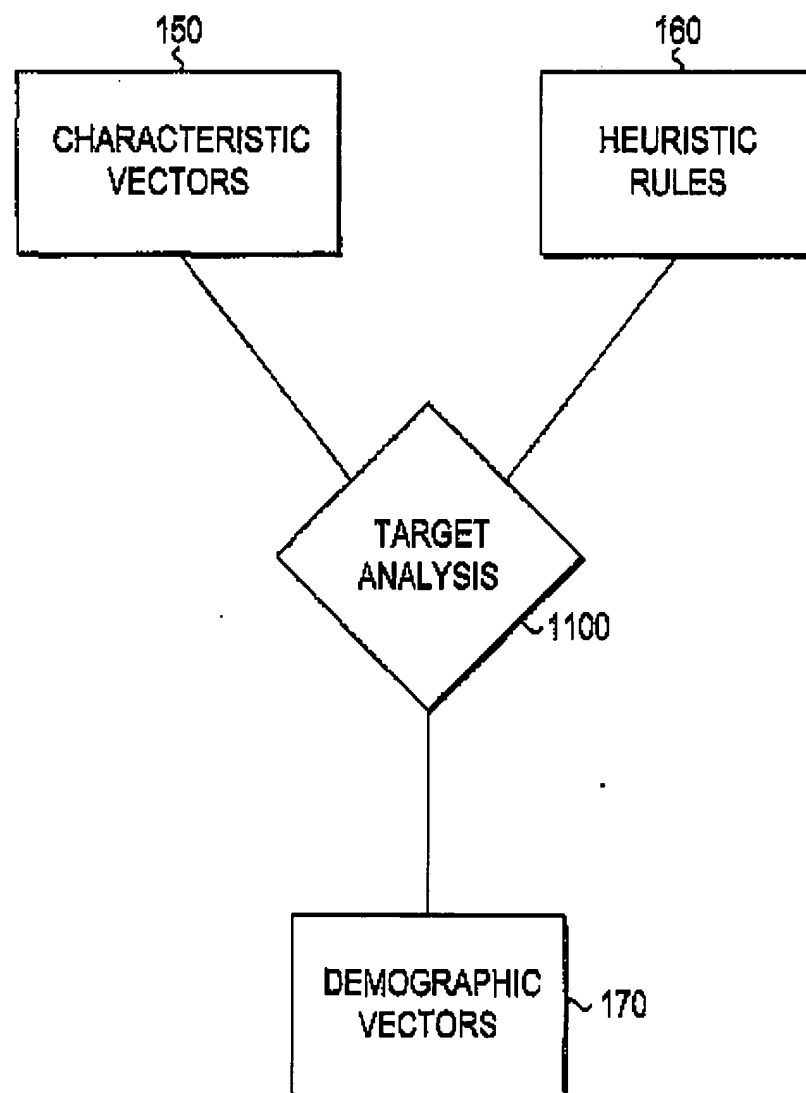
FIG. 10A

18/27

DEMOGRAPHIC GROUPS									
CATEGORIES	AGE			INCOME			SIZE		
	0-10	10-18	>70	0-20K	20-50K	50-100K	1	2	>5
NEWS	0.1	0.1	0.4	0.2	0.3	0.4	0.5	0.3	0.1
FICTION	0.5	0.3	0.2	0.4	0.2	0.3	0.3	0.2	0.1
FACTUAL	0.2	0.2	0.3	0.1	0.4	0.2	0.2	0.2	0.4
ADVERTISING	0.1	0.3	0.5	0.3	0.2	0.1	0.2	0.1	0.3
							0.3	0.7	
							0.8	0.2	
							0.4	0.6	
							0.5	0.5	

FIG. 10B

19/27

**FIG. 11**

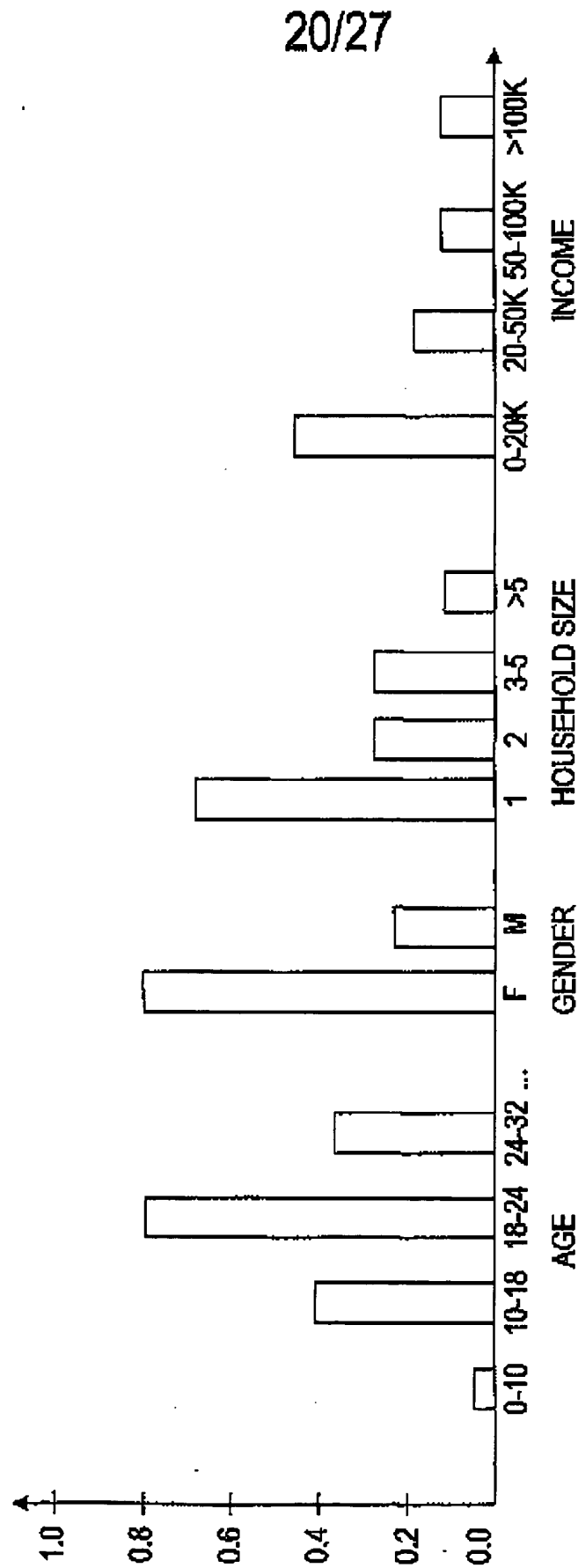
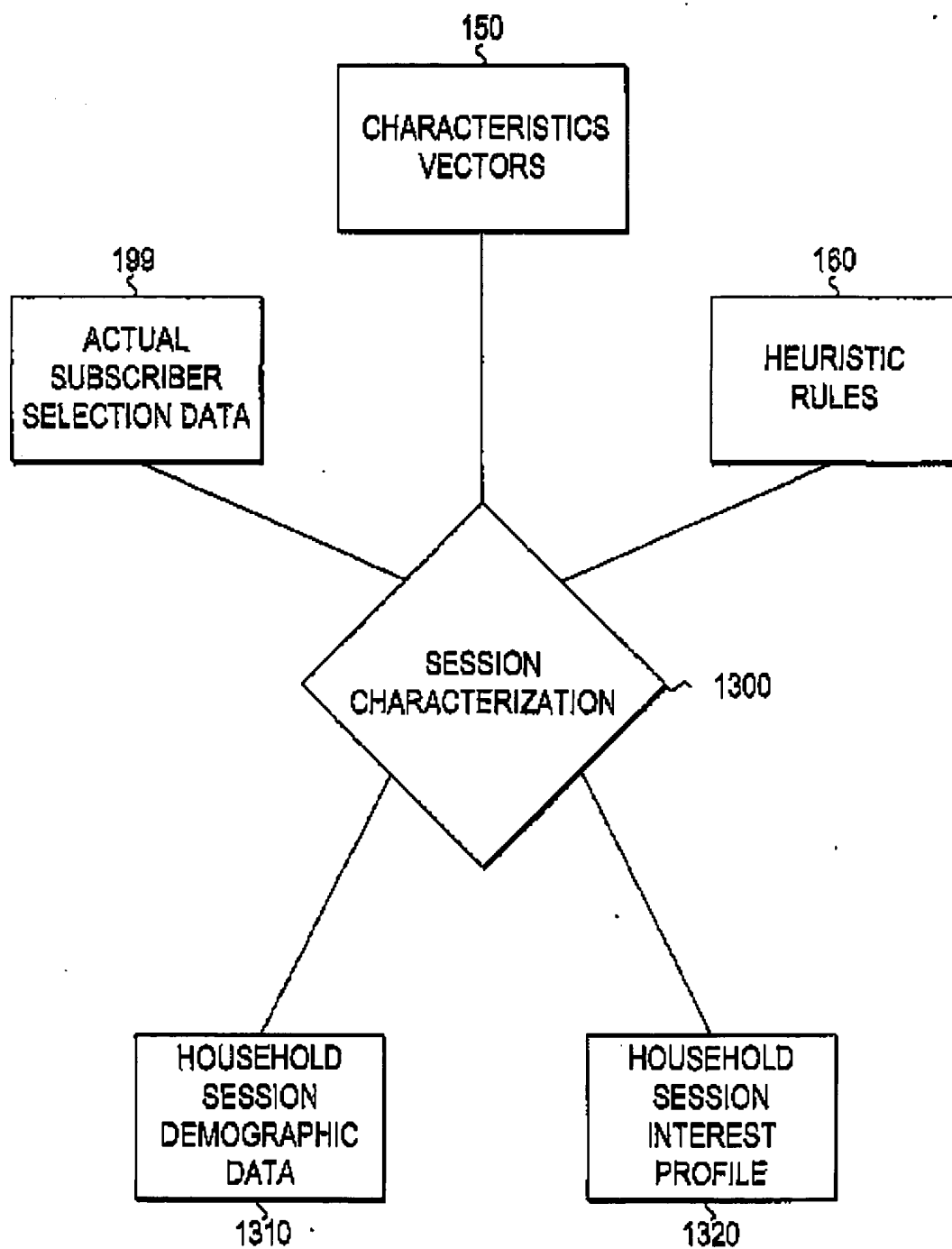
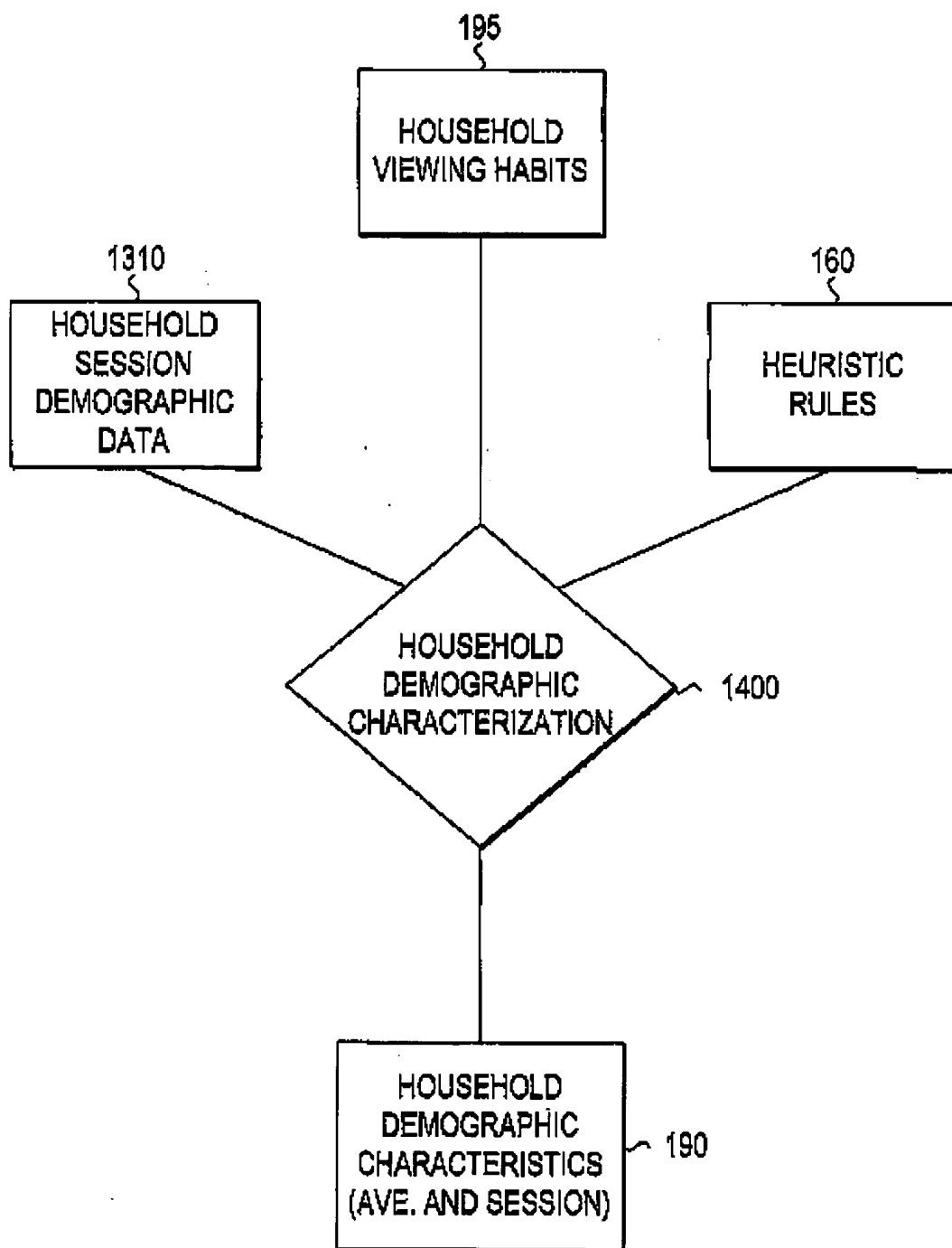


FIG. 12

21/27

**FIG. 13**

22/27

**FIG. 14**



23/27

1501 S	1505 S	1503 S	1507 S
HOUSEHOLD PARAMETER	AVERAGE VALUE	SESSION VALUE	UPDATE?
SIZE	2.6	3.0	YES
AGE	23.5	12	YES
SEX (FEMALE=1)	0.6	0.7	YES
INCOME (\$0-\$20K)	0.1	0.1	YES
INCOME (\$20-\$50K)	0.6	0.7	YES
INCOME (\$50-\$100K)	0.2	0.1	YES
INCOME (>\$100K)	0.1	0.1	YES
ZIP CODE			NO
TELEPHONE NUMBER			NO

FIG. 15

24/27

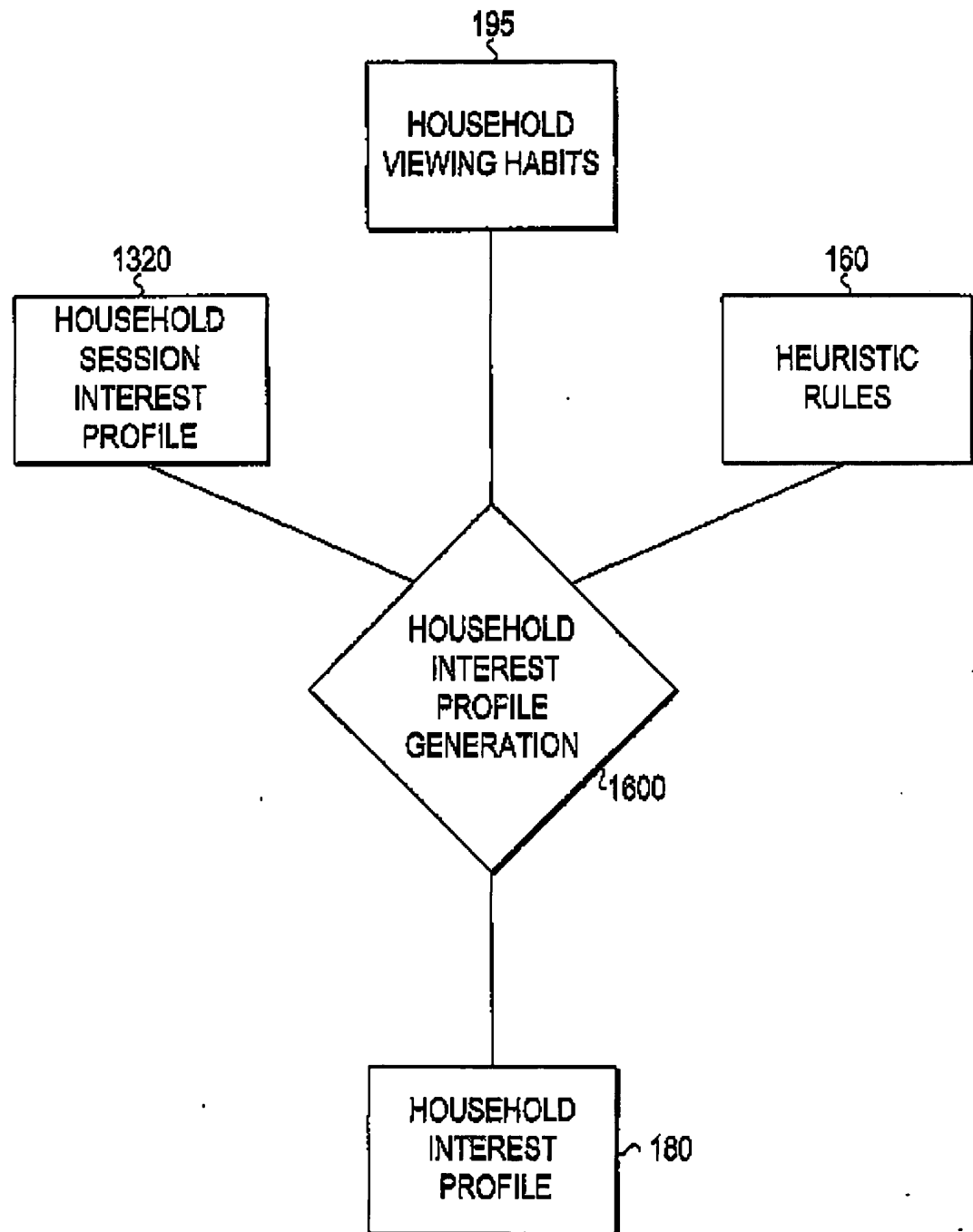


FIG. 16

25/27

		1701 HOUSEHOLD INTEREST	1703 AVERAGE VALUE	1705 SESSION VALUE
1709 PROGRAMMING	DRAMA		0.1	0.20
	ROMANCE		0.1	0.20
	ACTION		0.6	0.25
	SITCOM		0.2	0.30
	⋮			
	SPORTS		0	0.05
1707 PRODUCTS	HEALTH/EXERCISE		0.6	0.2
	FOOD		0.3	0.4
	CHILD-RELATED		0.0	0.1
	TOYS		0.0	0.1
	⋮			
	AUTOMOBILE		0.1	0.2

FIG. 17

26/27

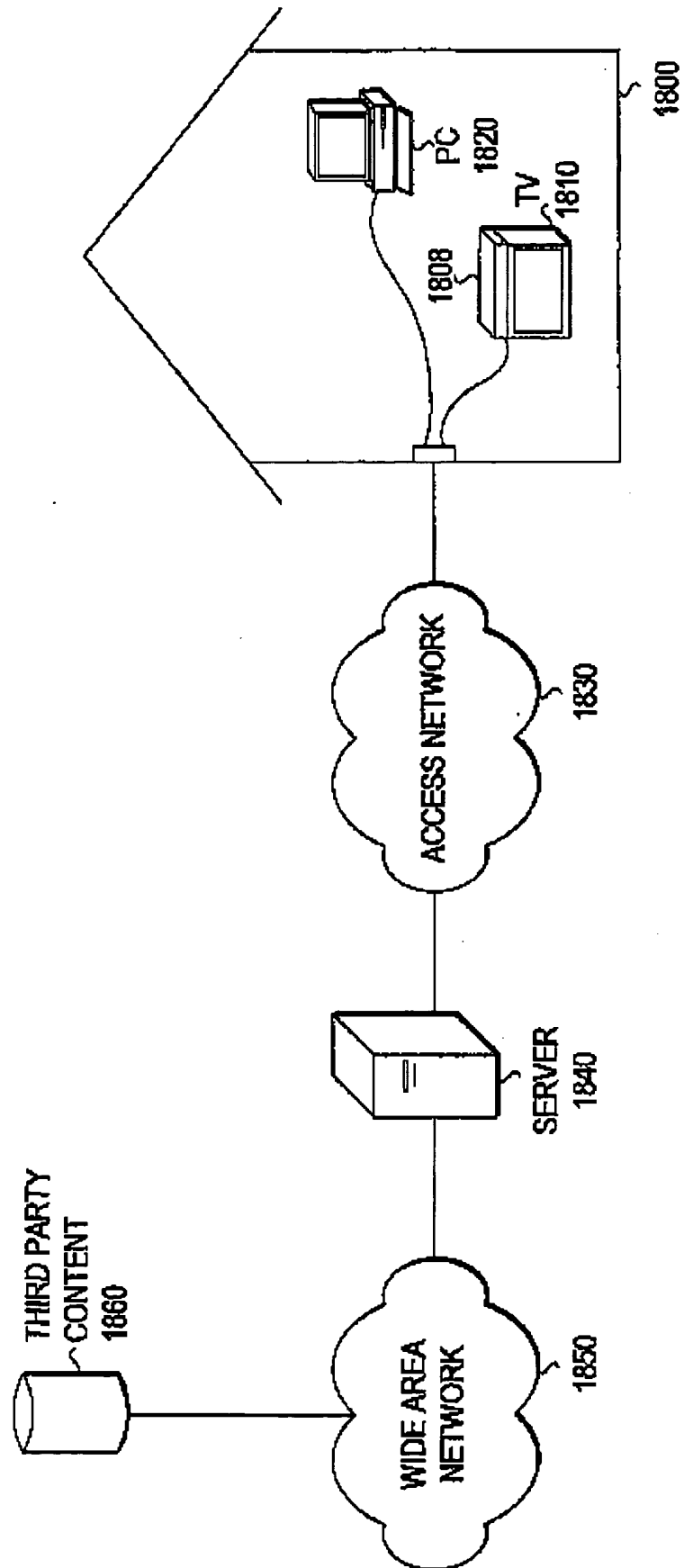


FIG. 18

27/27

1915	1921	1927	1933	1937
AD ID	PRODUCT	BRAND	% WATCHED	VOLUME
216	DIAPERS	HUGS	50%	6/10
1230	DETERGENT	SOAPY	90%	6/10
4137	AUTOMOBILES	SPEEDSTER	70%	8/10

**FIG. 19**

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/06459

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) : H04H 9/00; H04N 7/025, 7/10, 7/16

US CL : 725/14, 34

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 725/9, 14, 15, 16, 17, 18, 19, 20, 22, 32, 34, 36

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,758,257 A (HERZ et al) 26 May 1998, see whole document.	1-43
Y	US 5,774,170 A (HITE et al) 30 June 1998, see whole document.	1-43

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*B\* earlier document published on or after the international filing date

\*L\* document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\*

Inter document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\*

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\*

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*Z\*

document member of the same patent family

Date of the actual completion of the international search

01 JUNE 2001

Date of mailing of the international search report

17 JUL 2001

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks

Box PCT

Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JOHN W. MILLER

Telephone No. (703) 305-4795

Form PCT/ISA/210 (second sheet) (July 1998)\*